

IOWA STATE UNIVERSITY

Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and
Dissertations

1967

A numerical technique for solution of the linear second order elliptic equation in the plane

James Douglas Watson
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Mathematics Commons](#)

Recommended Citation

Watson, James Douglas, "A numerical technique for solution of the linear second order elliptic equation in the plane " (1967).
Retrospective Theses and Dissertations. 3438.
<https://lib.dr.iastate.edu/rtd/3438>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

This dissertation has been
microfilmed exactly as received 68-2870

WATSON, James Douglas, 1931-
A NUMERICAL TECHNIQUE FOR SOLUTION OF
THE LINEAR SECOND ORDER ELLIPTIC EQUATION
IN THE PLANE.

Iowa State University, Ph.D., 1967
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan

A NUMERICAL TECHNIQUE FOR SOLUTION OF THE LINEAR SECOND ORDER
ELLIPTIC EQUATION IN THE PLANE

by

James Douglas Watson

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Applied Mathematics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1967

TABLE OF CONTENTS

	Page
INTRODUCTION	1
DEVELOPMENT OF THE DIFFERENCE EQUATIONS	3
CONVERGENCE OF THE DIFFERENCE OPERATOR TO THE DIFFERENTIAL OPERATOR	29
EXISTENCE OF THE SOLUTION OF THE DIFFERENCE EQUATION AND CONVERGENCE TO THE SOLUTION OF THE DIFFERENTIAL EQUATION	46
COMPARISON OF METHODS AND RESULTS	59
BIBLIOGRAPHY	65
ACKNOWLEDGMENT	67
APPENDIX	68

INTRODUCTION

Consider the uniformly elliptic differential operator M defined by

$$(1.1) \quad \begin{aligned} M(u(x,y)) = & -(A(x,y)u(x,y))_{xx} - (B(x,y)u(x,y))_{xy} \\ & -(B(x,y)u(x,y))_{yx} - (C(x,y)u(x,y))_{yy} + D(x,y)u(x,y)_{xx} \\ & + E(x,y)u(x,y)_{yy} \end{aligned}$$

where A , B , and $C \in C^{1+\alpha}$, D and $E \in C^{\alpha}$, for $0 < \alpha < 1$. Since M is uniformly elliptic, there is a positive constant γ such that A and $C \geq \gamma$. Along with (1.1) we consider operator

$$(1.2) \quad Lu = Mu + Fu.$$

We are interested in solving, by finite difference techniques, the boundary value problem

$$(1.3) \quad M(u(x,y)) + F(x,y)u(x,y) = S(x,y) \text{ in } G$$

$$(1.4) \quad u(x,y) = g(x,y) \text{ on } G',$$

where G is an open connected region with piecewise C^1 boundary G' . Further assume $F \geq 0$, and F , S , and $g \in C^{\alpha}$. From Bers, John and Schechter (2, p. 136), we have strong enough conditions placed on the differential equation and the region so that a unique solution $u \in C^{2+\alpha}$ exists and depends continuously on $G + G'$ and the boundary condition g .

In this paper we propose a five-point difference scheme of positive type for solution of the differential equation (1.3). If the differential operator (1.1) is self-adjoint, then the corresponding difference operator will also be self-adjoint. With the conditions placed on the coefficients of the differential

equation the method is $o(1)$. If, instead, we assume A , B , and $C \in C^{2+\alpha}$ and D , E , F , and $S \in C^{1+\alpha}$, then the method is shown to be $O(h)$. It is possible for five-point difference schemes to be of order h^2 for certain Dirichlet problems. (See Mac Neal (11), or Varga (14, pp. 181-191), for example.) The proposed method is intended to be done numerically on a digital computer. In order to solve a general elliptic equation, this method cannot be shown to be of an order greater than h ; but in special cases this method may be $O(h^2)$.

Difference equations for solution of elliptic differential equations with a mixed derivative term have been proposed by Greenspan (6), (7), and Greenspan and Jain (8). Discussion of convergence of general difference schemes for elliptic differential equations has been done by Motzkin and Wasow (12), Bers (1), Greenspan and Parter (9) and others. Theorems on existence and uniqueness of the solution of finite difference equations and convergence of the solution of the difference equation to the solution of the differential equation have been done in the literature for difference equations of positive type. See, for instance, Parter (13) or Greenspan and Parter (9). Existence, uniqueness, and convergence proofs contained in this paper will be modifications of a somewhat more general theory to fit the needs of the problem at hand.

DEVELOPMENT OF THE DIFFERENCE EQUATIONS

The theoretical discussion of an arbitrary elliptic equation usually involves transforming it into some standard form, generally one involving the Laplacian and lower order derivatives in u . Weinberger (16, p. 44) or Garabedian (5, p. 59) shows that if a transformation

$$(2.1) \quad \xi = \xi(x, y), \quad \eta = \eta(x, y), \quad \xi, \eta \in C^2, \quad \frac{\partial(\xi, \eta)}{\partial(x, y)} \neq 0,$$

is performed on the differential operator (1.2), then

$$(2.2) \quad \begin{aligned} M(u) + Fu = & - \left[A \xi_x^2 + 2B \xi_x \xi_y + C \xi_y^2 \right] u_{\xi\xi} \\ & - \left[A \xi_x \eta_x + B \xi_x \eta_y + B \xi_y \eta_x + C \xi_y \eta_y \right] u_{\xi\eta} \\ & - \left[A \eta_x^2 + 2B \eta_x \eta_y + C \eta_y^2 \right] u_{\eta\eta} + M(\xi) u_\xi \\ & + M(\eta) u_\eta + Fu. \end{aligned}$$

Definition 2.1

We define a first canonical form of $Mu + Fu$ as

$$(2.3) \quad - \nabla \cdot \sigma(\xi, \eta) \nabla u + d(\xi, \eta) u_\xi + e(\xi, \eta) u_\eta + f(\xi, \eta) u.$$

This requires that the coefficient of $u_{\xi\eta}$ vanish and the coefficient of $u_{\xi\xi}$ be the same as that of $u_{\eta\eta}$. These conditions determine σ , d , e , and f , but not uniquely. Usually this is the desired form.

Definition 2.2

We define a second canonical form of $Mu + Fu$ as

$$(2.4) \quad \begin{aligned} & -(A^*(\xi, \eta) u_\xi)_\xi - (C^*(\xi, \eta) u_\eta)_\eta + D^*(\xi, \eta) u_\xi \\ & + E^*(\xi, \eta) u_\eta + F^*(\xi, \eta) u. \end{aligned}$$

Here the only requirement on the transformation is that the coefficient of $u_{\xi\eta}$ vanish. If the coefficient of $u_{\xi\eta}$ vanishes, then

$$(2.5) \quad A\xi_x\eta_x + B\xi_x\eta_y + B\xi_y\eta_x + C\xi_y\eta_y = 0.$$

Therefore, if $\xi_y \neq 0$,

$$(2.6) \quad \frac{A\xi_x\eta_x + B\xi_x\eta_y}{\xi_y} = -(B\eta_x + C\eta_y).$$

Hence, when $A\eta_x + B\eta_y \neq 0$

$$(2.7) \quad -\frac{\xi_x}{\xi_y} = \frac{dy}{dx} = \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y}.$$

If we choose $\eta \in C^2$, so that $A\eta_x + B\eta_y \neq 0$, we can solve for

ξ using (2.7). Then if $\xi = \text{constant}$ along solutions of

$$(2.8) \quad \frac{dy}{dx} = \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y},$$

the coefficient of the mixed derivative vanishes. Since the elliptic equation remains elliptic under the non-singular transformation (2.1), the coefficients of $u_{\xi\xi}$ and $u_{\eta\eta}$ have the same algebraic sign. This choice of ξ and η takes (1.2) into the second canonical form (2.4).

Techniques to obtain a first canonical form of (1.2) are discussed in Courant and Hilbert (4, pp. 154-163) or Garabedian (5, pp. 57-68).

If a transformation of (1.2) into either the first or second canonical form can be obtained, it might be most desirable to obtain the transformed equations and to proceed analytically as far

as possible. For a system in first or second canonical form, numerical procedures have been suggested by Varga (14, pp. 181-191), and others. Also, the numerical method we describe in this paper can be applied to the transformed equation with possibly a better rate of convergence than would be obtained on the original equation.

We assume the differential operator (1.2) cannot be put into a canonical form, and we must proceed numerically. The operator (1.2) is transformed into ξ - η coordinates, using the transformation which puts (1.2) into the second canonical form.

If we choose $\eta = \eta(x) = x$, then equation (2.8) becomes

$$(2.9) \quad \frac{dy}{dx} = \frac{B(x,y)}{A(x,y)}.$$

Then

$$\frac{\partial (\xi, \eta)}{\partial (x, y)} = \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} = -\xi_y.$$

Since

$$-\frac{\xi_x}{\xi_y} = \frac{dy}{dx} = \frac{B}{A},$$

and $A \geq c > 0$, dy/dx is bounded; hence $\xi_y \neq 0$ and $\frac{\partial (\xi, \eta)}{\partial (x, y)} \neq 0$.

We form a net on the region $G + G'$. Choose a horizontal and a vertical base line. On the horizontal line mark off intervals of length h_1 so that every point (x, y) in the region $G + G'$ has an x coordinate that falls in some interval of the base line.

The values of h_1 need not be equal; however, a uniform step size is convenient. Along the vertical line mark off intervals of width k_j , choosing the k_j all equal. After the net is computed, it might

appear that a different vertical spacing would be better, due to too large net spacing in some parts of $G + G'$. In that case, a different spacing might be chosen to correct this problem.

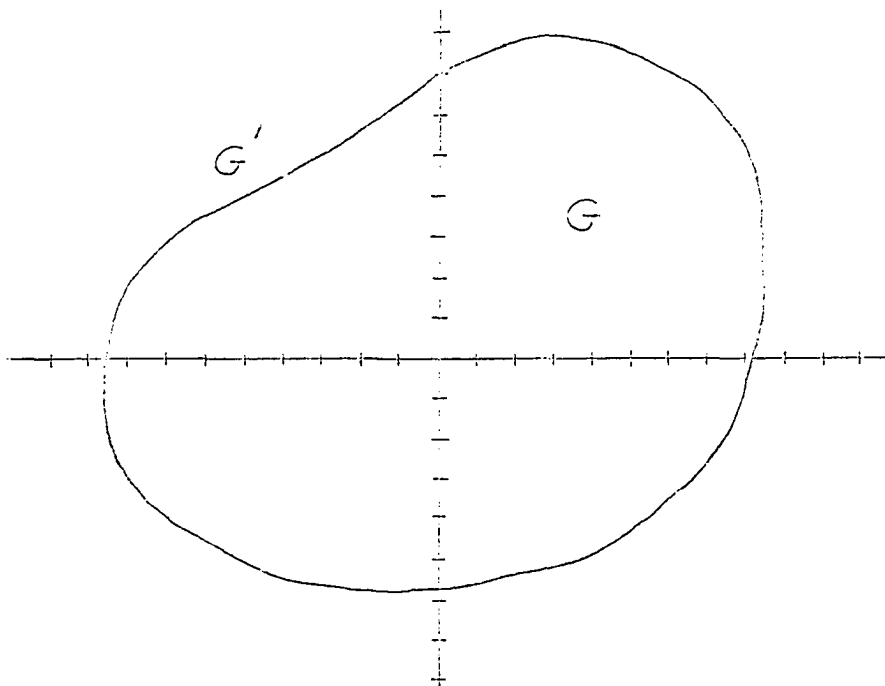


Figure 1. Region $G + G'$

Using each point on the vertical as a starting point, solve the differential equation $dy/dx = B/A$. Extend the solution in each direction until the solutions pass out of the region G . It may be necessary to add extra starting points on the vertical to cover the entire region $G + G'$ by the net. We now have a set of net points which are aligned vertically. $\eta = \text{constant}$ on net points on any vertical line. $\xi = \text{constant}$ along "curves" obtained by solving $dy/dx = B/A$.

We form the boundary as follows: Along the lines $\eta = \text{constant}$ connect the last interior point in $G + G'$ with the first point

outside. If, for example, the region is convex, then this happens at the top and bottom of the region. The intersection of the vertical line $\eta = \text{constant}$ with G' is taken as a boundary net point. No points outside $G + G'$ are used as net points. Along "curves" $\xi = \text{constant}$, connect the last point inside $G + G'$ with the first point outside by a straight line. The intersection of the line with G' is taken as a boundary net point.

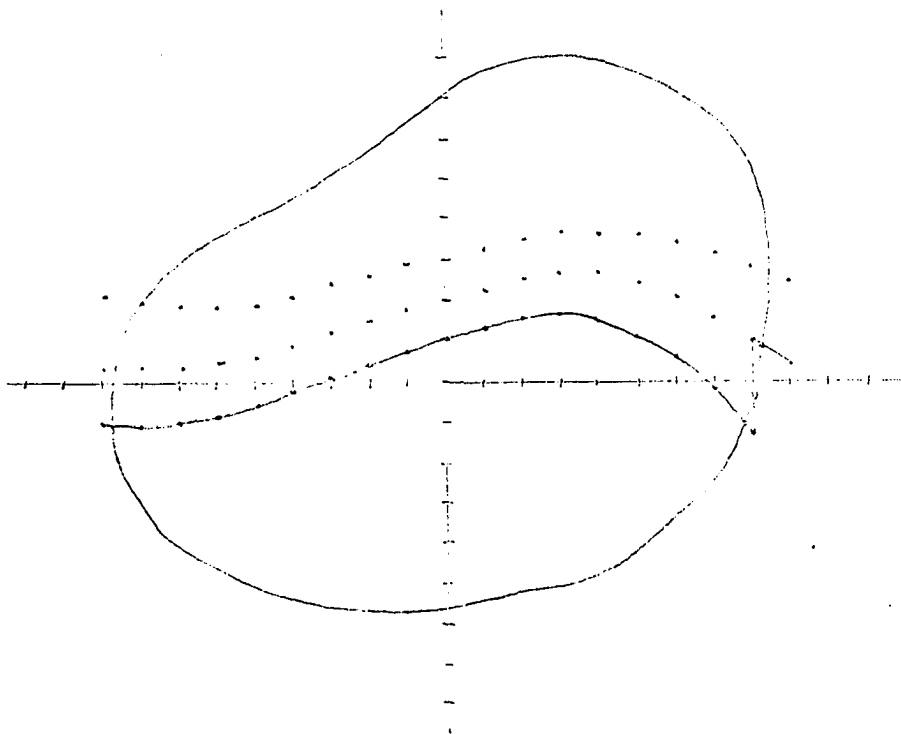


Figure 2. Net Points on $G + G'$

We require that the numerical method used to compute the net be of order h^2 . This is meaningful since first partials of B/A are C^∞ in $G + G'$. If higher order derivatives exist, then a higher order method may be practical to give faster convergence of the solution of the partial differential equation since the order

of the method depends upon the accuracy of the numerical procedure used to compute the net points. In case the differential operator is already in first or second canonical form, i.e. the mixed derivative is absent, the net is rectangular and it is not necessary to use the differential equation to form it.

Does the method described actually form the desired net? Do the solutions pass out of the region as stated? Do distinct starting points for the solutions of $dy/dx = B/A$, (2.9), remain distinct throughout the region? We shall examine these questions.

Since (1.1) is uniformly elliptic, A and $C \geq \delta > 0$ and $B/A \in C^1$. The partials are continuous and bounded on $G + G'$. Therefore, a local solution of $dy/dx = B/A$, (2.9), exists at every point in G . Two adjacent solutions cannot join since this would deny the uniqueness of the solution of (2.9) at the point the solutions meet.

We assert that the graph of every solution $y(x)$ of (2.9) can be extended from boundary to boundary of $G + G'$. Suppose the graph of $y(x)$ lies entirely within G . Since B/A is bounded away from zero on $G + G'$, the slope of the graph remains bounded. The local existence theorem assures us that the solution exists for some interval $\alpha_1 < x < \alpha_2$ about the initial starting value. Let us extend this solution to a maximum interval $\beta_1 < x < \beta_2$. However, by hypothesis we know a local solution exists at all points of $G + G'$; in particular, at the end points of the arc. In fact, by continuity of the solution of (2.9), the solution at the end points must be $y(x)$. Now the local existence theorem says the solution starting

at $(\beta_1, y(\beta_1))$ exists in an open interval containing β_1 ; hence β_1 is not an end point. This denies the fact that β_1 was an end point of the maximum interval. Consequently, the solutions of (2.9) must extend from boundary to boundary.

For each point in G we assign a neighborhood set $N(P)$. $N(P)$ and P are points to be used in the difference equation in which P is the center point. For our purposes $N(P)$ are the four adjacent points to P . We assume $P \notin N(P)$. Let $h > 0$ be a parameter which measures the size of neighborhoods (or the size of the net); specifically, let

$$h = \left\{ \max_{Q \in N(P)} |P - Q| \right\}$$

over all points P in the given net. Let $G(h)$ denote net points in G and $G'(h)$ the net points on G' .

Definition 2.3

For a net $G(h) + G'(h)$ we define connectedness as follows: Given two points $P \in G(h)$, $Q \in G(h) + G'(h)$, there is a "path" $P = P_1, P_2, \dots, P_r = Q$ from P to Q in which $P_{s+1} \in N(P_s)$. Note that since $N(P)$ is defined for interior points, all points P_i , $i = 1, r-1$ necessarily are in $G(h)$.

Since we have assumed that $G + G'$ is connected, we will require it of $G(h) + G'(h)$. If the net spacing is sufficiently small, this can be accomplished.

Let

$$(2.10) \quad \vec{t}_1 = (-\eta_y, \eta_x) / \sqrt{\eta_x^2 + \eta_y^2}$$

denote the tangent and the positive direction along the curve

$\eta = \text{constant}$. Let

$$(2.11) \quad \vec{t}_2 = (-\xi_y, \xi_x) / \sqrt{\xi_x^2 + \xi_y^2}$$

denote the tangent and the positive direction along the curve

$$\xi = \text{constant}.$$

We now show that if we resolve our differential equation in terms of derivatives along the \vec{t}_1 and \vec{t}_2 directions, a simple five-point difference equation can be obtained, and the effects of the mixed derivatives are gone. This part of the theory is carried out in full generality where η need not be a function of x alone. Therefore, other curvilinear nets could be used, rather than the one we have introduced and the corresponding difference equations obtained for that net.

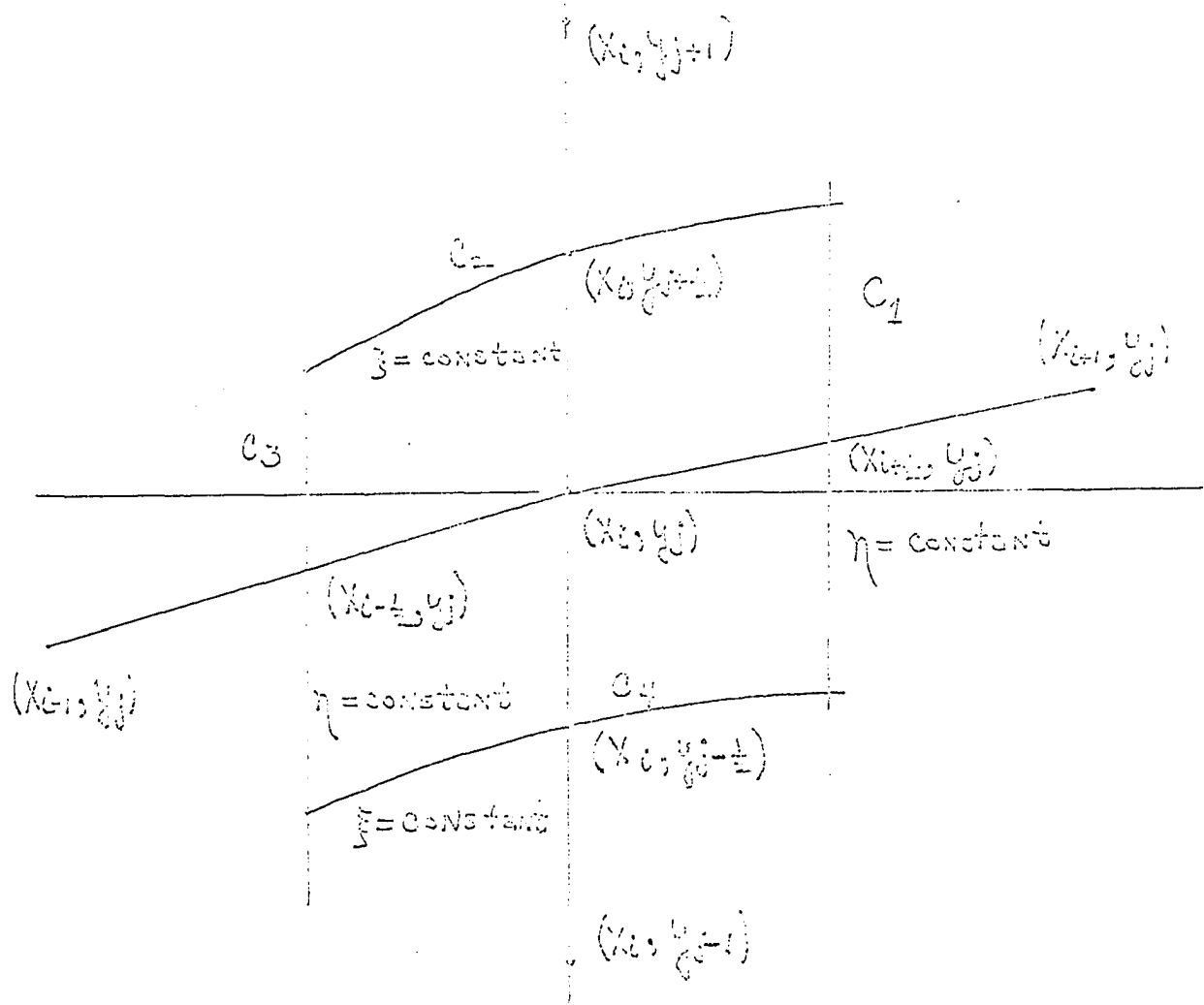
Following the technique in Varga (14, pp. 166-188), we form with each net point a net region. We then integrate the differential equation over this net region. This method of forming difference equations has the disadvantage that it has a low order discretization error. It possesses two possible advantages: First, if we are working with a self-adjoint differential equation, then the associated algebraic system is also self-adjoint. In studies in which eigenvalues are of interest, this might be quite important. The method can be extended to cases where the smoothness of the coefficients of (1.2) is only piecewise. Wachspress (15, p. 32-38 and p. 70-74) and Varga (14, p. 173-177, 190) discuss this for an equation in first canonical form.

For actual computation, the points $\{P_n\}$ are numbered

serially and some way is used for designating the points on the boundary. For the theoretical discussion, double subscripts are used for the points; so we have point $P_{i,j} = (x_i, y_j)$. The function evaluated at (x_i, y_j) is denoted by $u(x_i, y_j)$, $u(P_{i,j})$ or $u(i,j)$. When the function is evaluated halfway between two net points, it is denoted by $u(i + \frac{1}{2}, j)$.

Consider an arbitrary net point in $G(h)$ with coordinates (x_i, y_j) . As shown in Figure 3, the neighboring net points are (x_{i+1}, y_j) , (x_i, y_{j+1}) , (x_{i-1}, y_j) , and (x_i, y_{j-1}) . The mid-points on the lines connecting these points with (x_i, y_j) are $(x_{i+\frac{1}{2}}, y_j)$, $(x_i, y_{j+\frac{1}{2}})$, $(x_{i-\frac{1}{2}}, y_j)$, and $(x_i, y_{j-\frac{1}{2}})$ respectively.

We form a net region which contains the point $(x_i, y_j) \in G(h)$. Draw the curve $\eta = \text{constant}$ through $(x_{i+\frac{1}{2}}, y_j)$ and $(x_{i-\frac{1}{2}}, y_j)$. Then draw the curve $\xi = \text{constant}$ through $(x_i, y_{j+\frac{1}{2}})$ and $(x_i, y_{j-\frac{1}{2}})$. Extend the $\eta = \text{constant}$ curves and the $\xi = \text{constant}$ curves until the $\eta = \text{constant}$ and $\xi = \text{constant}$ curves intersect. This forms a region R associated with the point (x_i, y_j) . A net region R is formed for each point in $G(h)$. Figure 3 shows a possible region for $\eta = x$.

Figure 3. Net Region R

We define derivatives in the directions \vec{t}_1 and \vec{t}_2 and then integrate the differential equation over the net region R , writing it in terms of the new derivatives.

Definition 2.4

We define $u_{\vec{t}_1}$ to be the directional derivative in the direction of \vec{t}_1 ; thus $u_{\vec{t}_1} = \text{grad } u \cdot \vec{t}_1$. Similarly, $u_{\vec{t}_2} = \text{grad } u \cdot \vec{t}_2$.

Then, using (2.10) and (2.11)

$$(2.12) \quad u_{t_1} = (-\eta_y u_x + \eta_x u_y) / \sqrt{\eta_x^2 + \eta_y^2}$$

and

$$(2.13) \quad u_{t_2} = (-\xi_y u_x + \xi_x u_y) / \sqrt{\xi_x^2 + \xi_y^2}.$$

We solve for u_x and u_y in terms u_{t_1} and u_{t_2} .

$$(2.14) \quad u_x = (\xi_x \sqrt{\eta_x^2 + \eta_y^2} u_{t_1} - \eta_x \sqrt{\xi_x^2 + \xi_y^2} u_{t_2}) / (\eta_x \xi_y - \xi_x \eta_y)$$

and

$$(2.15) \quad u_y = (\xi_y \sqrt{\eta_x^2 + \eta_y^2} u_{t_1} - \eta_y \sqrt{\xi_x^2 + \xi_y^2} u_{t_2}) / (\eta_x \xi_y - \xi_x \eta_y).$$

To simplify writing, we let

$$(2.16) \quad z(\eta) = \sqrt{\eta_x^2 + \eta_y^2},$$

$$(2.17) \quad z(\xi) = \sqrt{\xi_x^2 + \xi_y^2},$$

and

$$(2.18) \quad J = \eta_x \xi_y - \xi_x \eta_y.$$

The differential equation is integrated over each net region R .

$$(2.19) \quad \iint_R [-(Au_x + Bu_y)_x - (Bu_x + Cu_y)_y + Du_x + Eu_y + Fu] dR = \iint_R SdR.$$

We use the divergence theorem to rewrite the first two terms of (2.19) in the form

$$(2.20) \quad \iint_R [-(Au_x + Bu_y)_x - (Bu_x + Cu_y)_y] dR \\ = - \int_C [(Au_x + Bu_y)n_x + (Bu_x + Cu_y)n_y] ds$$

where C is the boundary C_1 , C_2 , C_3 , and C_4 of the net region R .

(See Figure 3.) The derivatives on the right-hand side of (2.20) are written in terms of u_{t_1} and u_{t_2} using (2.14) and (2.15). Then

$$(2.21) \quad - \int_{C_1} [(Au_x + Bu_y)n_x + (Bu_x + Cu_y)n_y] ds \\ = \int_{C_1} \frac{(A\xi_x^2\eta_x + B\xi_y^2\eta_x + B\xi_x\eta_y + C\xi_y^2\eta_y)u_{t_1}}{J} ds \\ + \int_{C_1} \frac{z(\xi)}{J \cdot z(\eta)} (A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2)u_{t_2} ds \\ = \int_{C_1} \frac{z(\xi)}{J \cdot z(\eta)} (A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2)u_{t_2} ds,$$

since by (2.5) the coefficient of u_{t_1} vanishes. On C_3 the integral

has the same form except for a change of sign.

$$(2.22) \quad - \int_{C_3} [(Au_x + Bu_y)n_x + (Bu_x + Cu_y)n_y] ds$$

$$= - \int_{C_3} \frac{z(\xi)}{J \cdot z(\eta)} (A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2) u_{t_2} ds.$$

On C_2 (the top of the net region) we have

$$\begin{aligned} (2.23) \quad & - \int_{C_2} [(Au_x + Bu_y) n_x + (Bu_x + Cu_y) n_y] ds \\ & = \int_{C_2} \frac{z(\eta)}{J \cdot z(\xi)} (A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2) u_{t_1} ds \\ & = \int_{C_2} (A\eta_x\xi_x + B\eta_y\xi_x + B\eta_x\xi_y + C\eta_y\xi_y) u_{t_2} ds \\ & = \int_{C_2} \frac{z(\eta)}{J \cdot z(\xi)} (A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2) u_{t_1} ds \end{aligned}$$

since the coefficient of u_{t_2} vanishes. On C_4 the integral has the same form except for sign change.

$$\begin{aligned} (2.24) \quad & - \int_{C_4} [(Au_x + Bu_y) n_x + (Bu_x + Cu_y) n_y] ds \\ & = - \int_{C_4} \frac{z(\eta)}{J \cdot z(\xi)} (A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2) u_{t_1} ds. \end{aligned}$$

At this point we can see that if we discretize these equations we obtain a five-point difference formula where the effect of the mixed derivatives is absent. Note that along C_1 we need a difference analog of u_{t_2} which can be obtained from a difference quotient using $u(x_i, y_j) - u(x_{i+1}, y_j)$. Similarly on C_2 we need the discrete analog of u_{t_1} and this can be obtained from a difference quotient

using $u(x_i, y_{j+1}) - u(x_i, y_j)$. The vanishing of the coefficient of u_{t_1} on C_1 and the vanishing of the coefficient of u_{t_2} on C_2 makes this possible.

We will now return to our special net where $\eta = x$. Numerical integration does not give $\frac{y}{x}$; however, the integrals (2.21), (2.22), (2.23) and (2.24) can be written in terms of coefficients of the differential equation and the values of u in the net. Observe that

$$(2.25) \quad \eta_y = 0,$$

$$(2.26) \quad \frac{\xi_x}{\xi_y} = -B/A,$$

$$(2.27) \quad \vec{t}_1 = (0, 1),$$

$$(2.28) \quad \vec{t}_2 = \frac{(-A, B)}{\sqrt{A^2 + B^2}},$$

$$(2.29) \quad u_x = \frac{B}{A} u_{t_1} - \frac{\sqrt{A^2 + B^2}}{A} u_{t_2},$$

$$(2.30) \quad u_y = u_{t_1},$$

$$(2.31) \quad u_{t_1} = u_y,$$

and

$$(2.32) \quad u_{t_2} = -\frac{Au_x - Bu_y}{\sqrt{A^2 + B^2}}.$$

Then

$$(2.33) \quad \int_{C_1} \frac{z(\xi)}{J z(\eta)} (A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2) u_{t_2} ds \\ = \int_{C_1} \sqrt{A^2 + B^2} u_{t_2} ds$$

and

$$\begin{aligned}
 (2.34) \quad & - \int_{C_3} \frac{z(\eta)}{J \cdot z(\eta)} (A \eta_x^2 + 2B \eta_x \eta_y + C \eta_y^2) u_{t_2} ds \\
 & = - \int_{C_3} \sqrt{A^2 + B^2} u_{t_2} ds .
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 (2.35) \quad & - \int_{C_2} \frac{z(\xi)}{J \cdot z(\xi)} (A \xi_x^2 + 2B \xi_x \xi_y + C \xi_y^2) u_{t_1} ds \\
 & = - \int_{C_2} \frac{(AC - B^2)}{\sqrt{A^2 + B^2}} u_{t_1} ds ,
 \end{aligned}$$

and

$$\begin{aligned}
 (2.36) \quad & \int_{C_4} \frac{z(\eta)}{J \cdot z(\eta)} (A \eta_x^2 + 2B \eta_x \eta_y + C \eta_y^2) u_{t_1} ds \\
 & = \int_{C_4} \frac{(AC - B^2)}{\sqrt{A^2 + B^2}} u_{t_1} ds .
 \end{aligned}$$

Therefore, the line integral around the net region R ,

$$(2.20) \quad - \int_C [(Au_x + Bu_y) n_x + (Bu_x + Cu_y) n_y] ds$$

is equal to

$$\begin{aligned}
 (2.37) \quad & \int_{C_1} \sqrt{A^2 + B^2} u_{t_2} ds - \int_{C_2} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds \\
 & - \int_{C_3} \sqrt{A^2 + B^2} u_{t_2} ds + \int_{C_4} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds .
 \end{aligned}$$

Up to this point we have looked at only the first two terms of (2.19) integrated over the net region. The next two terms we rewrite as follows:

$$(2.38) \quad \iint_R Du_x dR = \iint_R D \left(\frac{-B}{A} u_{t_1} - \frac{\sqrt{A^2 + B^2}}{A} u_{t_2} \right) dR$$

and

$$(2.39) \quad \iint_R Eu_y dR = \iint_R Eu_{t_1} dR.$$

So the equation

$$(2.19) \quad - \int_C [(Au_x + Bu_y) n_x + (Bu_x + Cu_y) n_y] ds \\ + \iint_R (Du_x + Eu_y + Fu) dR = \iint_R S dR$$

is now written

$$(2.40) \quad \int_{C_1} \sqrt{A^2 + B^2} u_{t_2} ds - \int_{C_2} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds \\ - \int_{C_3} \sqrt{A^2 + B^2} u_{t_2} ds + \int_{C_4} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds \\ + \iint_R D \left(-\frac{B}{A} u_{t_1} - \frac{\sqrt{A^2 + B^2}}{A} u_{t_2} \right) dR \\ + \iint_R Eu_{t_1} dR + \iint_R F u dR = \iint_R S dR.$$

Up to this point no approximations have been made. We now obtain a set of equations which is a discrete approximation to equation (1.3) on $G + G'$. To carry this out, we first discretize the net region and then obtain an approximation to equation (2.40) on each of the net regions.

Consider an arbitrary net point with coordinates (x_i, y_j) . As shown in figure 4, the neighboring net points are (x_{i+1}, y_j) , (x_{i-1}, y_j) and (x_i, y_{j+1}) . The mid-points on the lines

connecting these points with (x_i, y_j) are $(x_{i+\frac{1}{2}}, y_j)$, $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$, $(x_{i-\frac{1}{2}}, y_j)$, and $(x_i, y_{j-\frac{1}{2}})$ respectively. To further simplify labeling, we label the central point 0, the four neighbor net points 1, 2, 3, 4, respectively, and the mid-points as 5, 6, 7, and 8, as shown in figure 4.

To form the net region, draw a line segment through point 5 with direction $\vec{t}_1(5)$. (\vec{t}_1 evaluated at point 5.) Make a similar construction at each of the points 6, 7 and 8. Extend these line segments in either direction until a quadrilateral is formed. The sides are labeled C_1 , C_2 , C_3 , and C_4 . This is the discrete net region for point 0. We refer to it as the "net region" in discussing the discretized equations. Figure 4 shows the region when $\eta = x$.

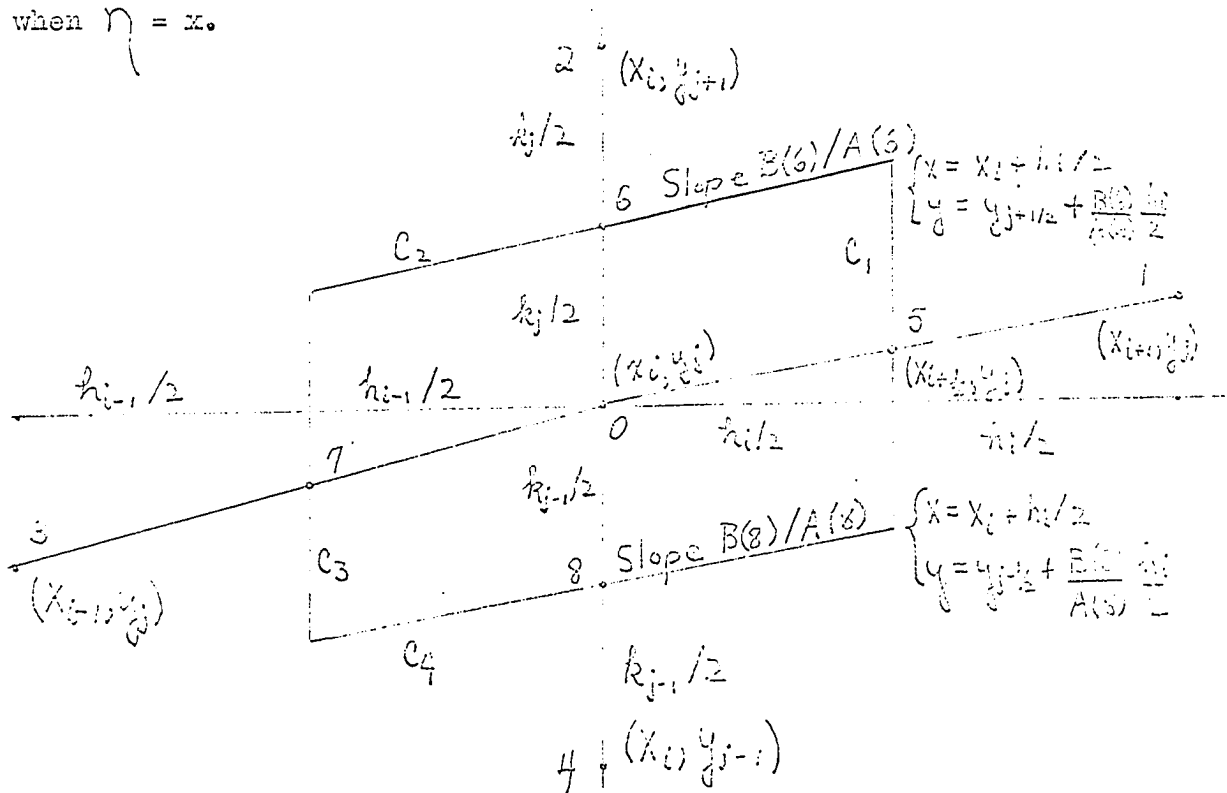


Figure 4. Typical Discretized Net Region

We now approximate equation (2.37) on the discretized net region.

On C_1 the function in the integrand of equation (2.33) is evaluated at 5. The derivative is replaced by the difference quotient evaluated at 5 and these constants are then taken outside of the integral. The remaining integral is replaced by the path length. Analogous treatment is given to the integrals on sides C_2 , C_3 and C_4 . This is done in equations (2.33), (2.34), (2.35) and (2.36).

In writing the difference quotient, the distance from point 1 to point 0 must be calculated. Rather than use the actual coordinates of point 1 and point 0 to compute the distance, the approxi-

mation $h_i \sqrt{1 + \frac{B^2(5)}{A^2(5)}}$ is used. This assumes the line from point 0

to point 1 has slope $B(5)/A(5)$ and this may not be exactly correct. Due to the continuity of B/A this error is small on a sufficiently small net region.

Along C_1

$$(2.41) \quad u_{t_2} \approx \frac{u(0) - u(1)}{h_i \sqrt{1 + (B(5)/A(5))^2}}$$

and the length of C_1 is

$$(2.42) \quad \frac{k_j + k_{j-1}}{2} + \frac{B(6)}{A(6)} \frac{h_i}{2} - \frac{B(8)}{A(8)} \frac{h_i}{2} ;$$

therefore,

$$(2.43) \quad \int_{C_1} \sqrt{A^2 + B^2} u_{t_2} ds \approx$$

$$\approx \frac{A(5)}{2} (u(0) - u(1)) \left\{ \frac{k_j + k_{j-1}}{h_i} + \frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \right\}.$$

Similarly

$$(2.44) \quad - \int_{C_3} \sqrt{A^2 + B^2} u_{t_2} ds \\ \approx - \frac{A(7)}{2} (u(3) - u(0)) \left\{ \frac{k_j + k_{j-1}}{h_{i-1}} - \frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \right\}.$$

Along C_2

$$(2.45) \quad u_{t_2} \approx \frac{u(2) - u(0)}{k_j}$$

and the length of C_2 is

$$\frac{h_i + h_{i-1}}{2} \sqrt{1 + (B(6)/A(6))^2};$$

therefore,

$$(2.46) \quad \int_{C_2} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds \\ \approx - \frac{h_i + h_{i-1}}{2k_j} \left\{ C(6) - \frac{B^2(6)}{A(6)} \right\} (u(2) - u(0)).$$

Similarly

$$(2.47) \quad - \int_{C_4} \frac{AC - B^2}{\sqrt{A^2 + B^2}} u_{t_1} ds \\ \approx \frac{h_i + h_{i-1}}{2k_{j-1}} \left\{ C(8) - \frac{B^2(8)}{A(8)} \right\} (u(0) - u(4)).$$

The integrals

$$(2.48) \quad \iint_R SdR,$$

$$(2.49) \quad \iint_R F u dR,$$

$$(2.39) \quad \iint_R E u y dR = \iint_R E u_{t_1} dR,$$

and

$$(2.38) \quad \iint_R D u_x dR = \iint_R D \left(-\frac{B}{A} u_{t_1} + \frac{\sqrt{A^2 + B^2}}{A} u_{t_2} \right) dR$$

remain to be discretized. In each case, the integrand is evaluated at the center of the net region and derivatives are replaced by forward differences. The net region R is a trapezoid whose area is

$$(2.50) \quad \frac{h_i + h_{i-1}}{2} \left\{ \frac{k_j + k_{j-1}}{2} + \frac{h_i - h_{i-1}}{2} \left(\frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \right) \right\}.$$

The area of the trapezoid is

$$(2.51) \quad \left(\frac{h_i + h_{i-1}}{2} \right) \cdot \left(\frac{k_j + k_{j-1}}{2} \right) + \text{a correction term of order } h^3$$

since

$$(2.52) \quad \frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \approx \frac{\partial}{\partial y} \left(\frac{B}{A} \right) (0) \cdot \frac{k_j + k_{j-1}}{2}.$$

We drop the correction term and use

$$(2.53) \quad a(i, j) = \left(\frac{h_i + h_{i-1}}{2} \right) \cdot \left(\frac{k_j + k_{j-1}}{2} \right) \text{ as the area of the}$$

net region R .

Now let us turn our attention to the integrals (2.48), (2.49), (2.39) and (2.38). The approximations are as follows:

$$(2.54) \quad \iint_R S dR \approx S(0) \cdot a(i, j) = S(0) \left(\frac{k_j + k_{j-1}}{2} \right) \cdot \left(\frac{h_i + h_{i-1}}{2} \right),$$

$$(2.55) \quad \iint_R F u dR \approx F(0) \cdot u(0) \cdot a(i, j) = F(0) u(0) \left(\frac{k_j + k_{j-1}}{2} \right) \left(\frac{h_i + h_{i-1}}{2} \right),$$

$$(2.56) \quad \iint_R E u_y dR =$$

$$\iint_R E u_{t_1} dR \approx E(0) \cdot \left(\frac{u(2) - u(0)}{k_j} \right) \cdot \left(\frac{h_i + h_{i-1}}{2} \right) \cdot \left(\frac{k_j + k_{j-1}}{2} \right),$$

and

$$(2.57) \quad \iint_R D u_x dR = \iint_R D \left(-\frac{B}{A} u_{t_1} - \frac{\sqrt{A^2 + B^2}}{A} u_{t_2} \right) dR$$

$$\approx D(0) \cdot \left(-\frac{B(0)}{A(0)} \frac{u(2) - u(0)}{k_j} + \frac{u(2) - u(0)}{h_i} \right) \cdot \left(\frac{h_i + h_{i-1}}{2} \right) \cdot \left(\frac{k_j + k_{j-1}}{2} \right).$$

Note that in (2.57) we used $B(0)/A(0)$ for the slope of the line from 0 to 1. This is not as good an approximation as $B(5)/A(5)$ but gives a simpler discretization and does not increase the order of the discretization error.

When we make use of the above results, the discretization of (2.40) on the net region R gives

$$(2.58) \quad \lambda(i, j) u(i, j) - \lambda(i, j) u(i-1, j) - r(i, j) u(i+1, j)$$

$$- \tau(i, j) u(i, j+1) - b(i, j) u(i, j-1)$$

$$= s(i, j) + a(i, j) \hat{c}(i, j),$$

where $\hat{c}(i, j)$ is the truncation error which absorbs all the errors of discretization, and where

$$(2.59) \quad \lambda(i, j) = \frac{A(i-\frac{1}{2}, j)}{2h_i} \frac{k_j + k_{j-1}}{2} + \frac{B(i, j-\frac{1}{2})}{A(i, j-\frac{1}{2})} - \frac{B(i, j+\frac{1}{2})}{A(i, j+\frac{1}{2})},$$

$$(2.60) \quad r(i, j) = \frac{A(i+\frac{1}{2}, j)}{2h_{i-1}} \frac{k_j + k_{j-1}}{2} - \frac{B(i, j-\frac{1}{2})}{A(i, j-\frac{1}{2})} + \frac{B(i, j+\frac{1}{2})}{A(i, j+\frac{1}{2})} +$$

$$+ D(i,j) \frac{a(i,j)}{h_i} ,$$

$$(2.61) \quad t(i,j) = \frac{1}{k_j} \left[\left\{ C(i,j+\frac{1}{2}) - \frac{B^2(i,j+\frac{1}{2})}{A(i,j+\frac{1}{2})} \right\} \frac{h_1+h_{j-1}}{2} - \frac{B(i,j)B(i,j)}{A(i,j)} \cdot a(i,j) + B(i,j)a(i,j) \right] ,$$

$$(2.62) \quad b(i,j) = \left\{ C(i,j-\frac{1}{2}) + \frac{B^2(i,j-\frac{1}{2})}{A(i,j-\frac{1}{2})} \right\} \frac{h_1+h_{j-1}}{2k_{j-1}} ,$$

$$(2.63) \quad d(i,j) = \ell(i,j) + r(i,j) + t(i,j) + b(i,j) + F(i,j)a(i,j),$$

and

$$(2.64) \quad s(i,j) = S(i,j)a(i,j) .$$

The $c(i,j)$ term is dropped and the resulting equation is the discretized equation used to compute u at the point (x_i, y_j) .

For each point (x_i, y_j) of the net where $u(i,j)$ is unknown, we write equation (2.58). Therefore, there is one linear equation derived at each point where the value of u is unknown. Each point is coupled to the four adjacent neighbors, giving rise to a "five-point" formula. Note also that the contribution to these equations from the second order terms of the differential equation does not change order of magnitude with order of magnitude change in step size. Therefore, the equations do not have to be rescaled with change in net size in machine computation. However, for comparison with the differential equation, the contributions to the algebraic equations must be divided by $a(i,j)$ and will change in magnitude as $h \rightarrow 0$.

In setting up the system of equations, we follow Varga (14, p. 167). First, number the net points where u is unknown along

$\}$ = constant from left to right and top to bottom, calling this the "natural ordering" of the net region as shown in figure 2.

Then number the points where u is known, say, on the boundary.

Transpose the parts of the equation that have been evaluated to the right-hand side. We then arrange each equation with the unknowns in ascending order. In writing this system, we put the equation of each net point below that of the previous net point. This gives a system of equations

$$(2.55) \quad A u = c.$$

A is the coefficient matrix of the unknowns, and u is the vector of unknown values $u(i,j)$ at net points. c is a vector which is formed from S evaluated at net points and from the solution function u given in boundary conditions of the problem.

Note the following properties of A :

1) A is real.

2) Since $A > 0$, $C > 0$, and $AC - E^2 > 0$, h can be taken small enough, depending on the magnitude of D , A , B , C and E so that

ℓ , r , t and $b > 0$ in the difference equation (2.50). In equations (2.55), (2.56) and (2.57) the coefficient of F is multiplied by an $O(h^2)$ factor, the coefficients of E and D are multiplied by an $O(h)$ factor, while the remaining terms of the difference equation do not approach zero as $h \rightarrow 0$.

3) For $P \in G(h)$ and $N(P) \in G(h)$, $d = \ell + r + t + b + F \cdot a(i,j)$.

Since $F \geq 0$ and all other terms are positives if the conditions of property 2 are satisfied, A has positive diagonal elements and non-positive off diagonal elements.

4) If $P \in G(h)$ and some $Q \in N(P)$ in G' , then d is greater than the sum of the off diagonal elements since one or more off diagonal terms has been transposed to the right hand side of the equation before A was formed and is not an unknown. This transposition of an off diagonal term does not affect the magnitude of the diagonal, but reduces the sum of off diagonal terms in the equation corresponding to point P .

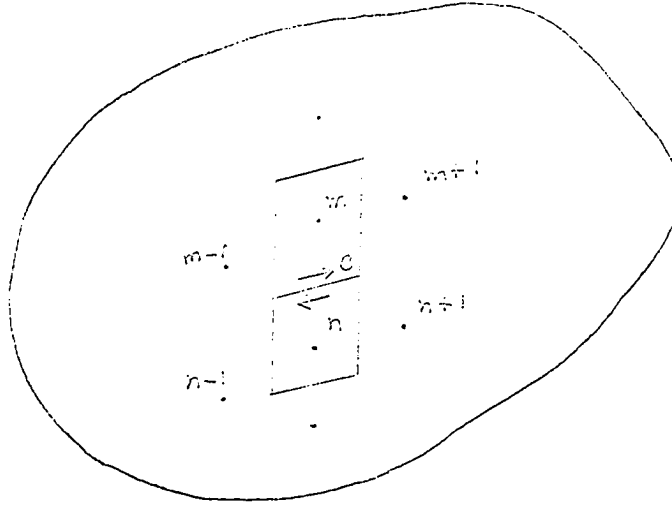
5) If $F > 0$, A is strictly diagonally dominant. If $F = 0$, strict dominance occurs only in rows of A where the corresponding net point has neighbors on G' .

6) If $D = E = 0$, then the differential operator L and the corresponding difference operator are both self-adjoint.

Proof of 6):

The self-adjointness of L_u is evident.

In matrix A there are as many rows as net points at which u is unknown in the region R . The m th row of A is the set of coefficients of the equation corresponding to the m th net point in the region. The ordering of the unknowns and equations is such that the diagonal entries of A are the elements d of (2.58). Let us examine the element α_{mn} of A . The index m corresponds to the m th net point and the index n to the n th net point in the serial ordering of the net points of R . That is, α_{mn} is the coefficient of the n th unknown in the m th equation. If net point m and net point n are not neighbors, this coefficient is zero.

Figure 5. Net Regions for Points m and n

Now let us take a specific example as shown in figure 5 where net point n lies directly below net point m along the curve $\eta =$ constant. The coefficient α_{mn} is computed by an approximation of a line integral of the form

$$(2.65) \quad \int_C f(x,y) u_{t_1} n_{t_2} ds$$

directed along the path from left to right. Our discrete approximation of this value is

$$(2.66) \quad f(x^*, y^*) \frac{u(m) - u(n)}{h} (-n_{t_2}) \cdot (\text{path length}),$$

where $f(x^*, y^*)$ is the evaluation of $f(x, y)$ at the mid-point between points m and n (see figure 5), h is the positive distance from m to n , and the normal is directed down (outward) as the minus sign on n_{t_2} indicates. The coefficient of $u(n)$ in (2.66) is

$$(2.67) \quad f(x^*, y^*)(-n_{t_2})(\text{path length}) / h.$$

This is the entry in the m th row and n th column of A .

We next look at the n th row of A . The term α_{nm} (the coefficient of $u(m)$) in this equation is derived from the line integral

$$(2.68) \quad \int_C f(x, y) u_{t_1} n_{t_2} ds.$$

The integration is from right to left; hence the outward normal is n_{t_2} . The discrete approximation of (2.68) is

$$(2.69) \quad f(x^*, y^*) \frac{u(n) - u(m)}{h} n_{t_2} \cdot (\text{path length}),$$

where the terms are defined as before. The coefficient of $u(m)$,

$$(2.70) \quad -f(x^*, y^*) n_{t_2} \cdot (\text{path length}) / h,$$

is identical to (2.67), i.e. $\alpha_{nm} = \alpha_{mn}$. This establishes the symmetry of A .

7) If $D = E = 0$, results of Varga (14, pp. 17-23) show that A is positive definite.

CONVERGENCE OF THE DIFFERENCE OPERATOR TO THE DIFFERENTIAL OPERATOR

We derived the difference equation

$$(3.1) \quad d(i,j)u(i,j) - r(i,j)u(i+1,j) - \lambda(i,j)u(i-1,j) \\ - t(i,j)u(i,j+1) - b(i,j)u(i,j-1) = s(i,j).$$

For computation purposes it is advantageous to have the operator in this form since the magnitude of the terms of this operator is not changed significantly by order of magnitude changes in net size h . Now consider the difference operator $L_h u$ obtained by dividing the left hand side of (3.1) by the approximate area $a(i,j)$.

$$(3.2) \quad L_h u = (d(i,j)u(i,j) - r(i,j)u(i+1,j) - \lambda(i,j)u(i-1,j) \\ - t(i,j)u(i,j+1) - b(i,j)u(i,j-1)) / a(i,j).$$

$L_h u$ is compared to Lu , or, alternatively, we work with $M_h u = L_h u - Fu$.

$$(3.3) \quad M_h u = M_1 u + M_2 u + M_3 u + M_4 u + M_5 u + M_6 u$$

where

$$(3.4) \quad M_1 u = \frac{2 A(5)}{h_i + h_{i-1}} \left\{ \frac{1}{h_i} + \frac{1}{k_j + k_{j-1}} \left(\frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \right) \right\} (u(0) - u(1)),$$

$$(3.5) \quad M_2 u = - \frac{2 A(7)}{h_i + h_{i-1}} \left\{ \frac{1}{h_i} - \frac{1}{k_j + k_{j-1}} \left(\frac{B(6)}{A(6)} - \frac{B(8)}{A(8)} \right) \right\} (u(3) - u(0)),$$

$$(3.6) \quad M_3 u = - \frac{2}{k_j(k_j + k_{j-1})} \left\{ c(6) - \frac{B^2(6)}{A(6)} \right\} (u(2) - u(0)),$$

$$(3.7) \quad M_4 u = + \frac{2}{k_{j-1}(k_j + k_{j-1})} \left\{ c(8) - \frac{B^2(8)}{A(8)} \right\} (u(0) - u(4)),$$

$$(3.8) \quad M_5 u = + D(0) \left\{ - \frac{B(0)}{A(0)} + \frac{u(2) - u(0)}{k_j} + \frac{u(1) - u(0)}{h_i} \right\},$$

and

$$(3.9) \quad M_6 u = + E(0) \frac{u(2) - u(0)}{k_j}.$$

Since $Lu = Mu + Fu$, we compare $M_h u$ with

$$(1.1) \quad Mu = - (Au_x + Bu_y)_x - (Bu_x + Cu_y)_y + Du_x + Eu_y,$$

where both Mu and $M_h u$ are evaluated at the same interior point we denote by O .

Two theorems can be proved. The first requires more smoothness of the coefficients of the differential equation. It establishes the "order of the method" or what is the maximum rate of convergence that can be expected of the method in the general case. The second uses only strong enough conditions on the coefficients of the differential equation to insure the solution of the differential equation to be a continuous function of the region and the boundary conditions.

Theorem 3.1a

Let Mu and $M_h u$ be the operators previously defined. Suppose $A, B, C \in C^{2+\alpha}$, A and C are bounded away from zero, and D and $E \in C^{1+\alpha}$. Let $u \in C^3$. Let the method used to compute the net region be order h^2 . Then there is a constant $K > 0$ and an h_0 such that when $h \leq h_0$

$$(3.10) \quad |Mu - M_h u| \leq K h$$

in $G(h)$.

Theorem 3.1b

Let Mu and $M_h u$ be the operators previously defined. Suppose $A, B, C \in C^{1+\alpha}$ and $D, E \in C^\alpha$. A and C are bounded away from zero. Let $u \in C^2$. Let the method used to compute the net region be order h^2 . Then

$$(3.11) \quad \lim_{h \rightarrow 0} \|Lu - L_h u\| = \lim_{h \rightarrow 0} \|Mu - M_h u\| = 0.$$

or, equivalently, there is a function $w(h)$ such that

$\lim_{h \rightarrow 0} w(h) = 0$, and

$$(3.12) \quad \|Mu - M_h u\| \leq w(h)$$

for all h sufficiently small.

These theorems are needed to prove convergence theorems. We show that the convergence of the discrete solution to the continuous solution is at a rate proportional to h when $\|M_h u - Mu\| \leq Kh$. This is the best that can be obtained with this five-point formula and the net formed as described. Only Theorem 3.1a is proved. Theorem 3.1b follows with appropriate modifications.

Before showing the convergence of the difference operator M_h to the differential operator M , we need some preliminary estimates.

Once the net is determined, the slope s of the line, say, from 0 to 1 is determined. We show that this slope s is closely approximated by $(B/A)(5)$, the slope of solution to $y' = B/A$ evaluated at the midpoint of the line from 0 to 1.

Corresponding to the two theorems 3.1a and 3.1b we get two sets of theorems. They have the a and b labels throughout this section. The first is the stronger theorem and necessarily has a stronger hypothesis.

Theorem 3.2a

Let A , B , and $C \in C^2$. Let A , $C \geq \gamma > 0$. Suppose the net region is calculated by a method of order h^2 . Let s be the slope

of the line from 0 to 1. Then there exist constants K and h_0 such that $h \leq h_0$ implies

$$(3.13) \quad \left| s - \frac{B}{A} (5) \right| \leq K h^2.$$

Theorem 3.2b

Let A , B , and $C \in C^1$. Let A , $C \geq \gamma > 0$. Suppose the net region is calculated by a method of order h^2 . Let s be the slope of the line from 0 to 1. Then there is a function w , with

$\lim_{h \rightarrow 0} w(h) = 0$, and h_0 such that $h \leq h_0$ implies

$$(3.14) \quad \left| s - \frac{B}{A} (5) \right| \leq h w(h).$$

Similar estimates hold at point 7.

Before proving Theorem 3.2a some preliminary results are obtained. We first quote two theorems from Coddington and Levinson (3, p. 22):

Theorem 7.1 - Let $f \in (C, \text{Lip})$ in a domain D of the $(n+1)$ dimensional (t, x) space, and suppose ψ is a solution of $x' = f(t, x)$ on an interval $I: a \leq t \leq b$. [then, if] there exists a $\delta > 0$ such that for any $(\tau, \xi) \in U$, where $U: a < \tau < b \quad |\xi - \psi(\tau)| < \delta$ there exists a unique solution ϕ of $x' = f(t, x)$ on I with $\phi(\tau, \tau, \xi) = \xi$. Moreover, $\phi \in C$ on the $(n+2)$ - dimensional set $V: a < t < b \quad (\tau, \xi) \in U$

For our problem $f = B/A$, D is two-dimensional (x, y) space.

We will be interested not in this theorem but its successor

(3, p. 25):

Theorem 7.2 - Let the hypothesis of 7.1 [above] be satisfied, and suppose f_x exists and $f_x \in C$ in D . Then $\phi \in C^1$ on V and moreover

$$\det \Phi_{\xi}(t, \tau, \xi) = \exp \int_{\tau}^t \operatorname{tr} f_x(s, \phi(s, \tau, \xi)) ds.$$

Note if $f(x, y) \in C^2$, then $y \in C^2$. Further $\frac{\partial^3 y}{\partial x^3}$ and $\frac{\partial^3 y}{\partial x^2 \partial \xi}$ are continuous. We use the continuity of $\frac{\partial^3 y}{\partial x^3}$ and $\frac{\partial^3 y}{\partial x^2 \partial \xi}$ in the proof of Theorem 3.2a.

From Henrici (10, p. 64) a general one-step method is written in the form

$$(3.15) \quad y_{n+1} - y_n = h \Phi(x_n, y_n; h)$$

where Φ is called the increment function. Henrici (10) gives a definition on pages 64-65 and 68 for the order of a "one-step method".

We summarize as follows:

Let (x_n, y_n) be in $a \leq x_0 \leq b$, $-\infty < y_0 < \infty$. Let $y(x, x_n, y_n)$ be a solution of $y' = f(x, y)$, $f \in C^1$, passing through (x_n, y_n) .

A method is called "of order p " if there are constants K and h_0 such that $h \leq h_0$ implies

$$(3.16) \quad y(x_n + h, x_n, y_n) - y(x_n, x_n, y_n) - h \Phi(x_n, y_n; h) \leq K h^{p+1},$$

where p is the largest integer for which this inequality is possible.

If the differential equation $y' = B/A$ is solved by a "one-step" method, we require that the order be at least 2. Then

$$(3.17) \quad \begin{aligned} & \left| y(x_{n+1}, x_n, y_n) - y_{n+1} \right| \\ &= \left| y(x_{n+1}, x_n, y_n) - y_n - h \Phi(x_n, y_n; h) \right| \end{aligned}$$

$$= |y(x_{n+1}, x_n, y_n) - y(x_n, x_n, y_n) - h \bar{Q}(x_n, y_n, h)| \leq K h^3$$

If a multi-step method is used to obtain the net region, we require it be such that the local discretization error in $G + G'$ be of the form

$$(3.18) \quad |y(x_{n+1}, x_n, y_n) - y_{n+1}| \leq K h^3$$

for suitable constant K and h sufficiently small. Regardless of the method used to obtain the net region, we refer to it as of order h^2 and it must satisfy (3.18).

Proof of Theorem 3.2a:

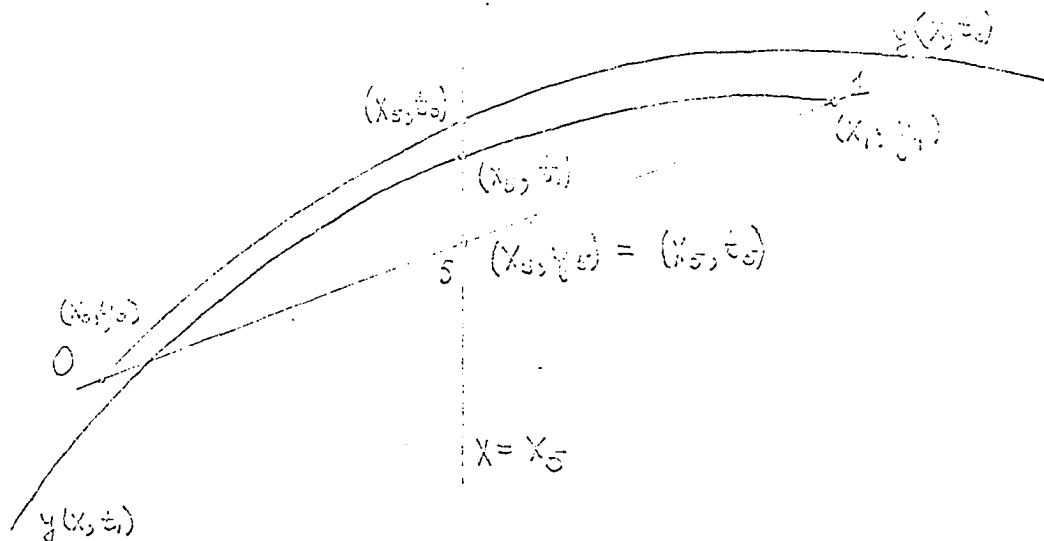


Figure 6. Typical Representation of Solutions of $dy/dx = B/A$ and a Straight Line Approximation

Let (x_1, y_1) be the coordinates of point 1, (x_5, y_5) the coordinates of point 5, and (x_0, y_0) the coordinates of point 0. Let t_0 be the intercept on the line $x = x_5$ of the solution of $y' = B/A$ passing through (x_0, y_0) . Denote this solution by $y(x, t_0)$. Let $y(x, t_1)$ and $y(x, t_5)$ be similarly defined. Let $y^*(x)$ be the solution through $(x_5, \frac{t_1 + t_0}{2})$.

The exact slope of the line from 0 to 1 is written

$$(3.19) \quad s = \frac{y(x_1, t_1) - y(x_0, t_0)}{h_i}.$$

The coordinate

$$(3.20) \quad y_5 = \frac{1}{2}(y(x_1, t_1) + y(x_0, t_0)).$$

We wish to compare s with $\frac{B}{A}(5)$. First we expand s about point 5.

$$(3.21) \quad s = \frac{1}{h_i} \left\{ y(5) + \frac{\partial y}{\partial x}(5)(x_1 - x_5) + \frac{\partial y}{\partial t}(5)(t_1 - t_5) \right\} \\ - \frac{1}{h_i} \left\{ y(5) + \frac{\partial y}{\partial x}(5)(x_0 - x_5) + \frac{\partial y}{\partial t}(5)(t_0 - t_5) \right\} \\ + \frac{1}{h_i} \left\{ \frac{1}{2} \frac{\partial^2 y}{\partial x^2} (x_1 - x_5)^2 + \frac{\partial^2 y}{\partial x \partial t} (x_1 - x_5)(t_1 - t_5) \right. \\ \left. + \frac{1}{2} \frac{\partial^2 y}{\partial t^2} (t_1 - t_5)^2 \right\} (Q_1) - \frac{1}{h_i} \left\{ \frac{1}{2} \frac{\partial^2 y}{\partial x^2} (x_0 - x_5)^2 \right. \\ \left. + \frac{\partial^2 y}{\partial x \partial t} (x_0 - x_5)(t_0 - t_5) + \frac{1}{2} \frac{\partial^2 y}{\partial t^2} (t_0 - t_5)^2 \right\} (Q_2),$$

where Q_1 and Q_2 are suitable mean values.

$$(3.22) \quad x_1 - x_5 = -(x_0 - x_5) = h_i/2.$$

Further rewrite

$$(3.23) \quad t_1 - t_5 \text{ as } (t_1 - t_0)/2 + [(t_1 + t_0)/2 - t_5]$$

and

$$(3.24) \quad t_0 - t_5 \text{ as } (t_0 - t_1)/2 + [(t_1 + t_0)/2 - t_5].$$

Then (3.21) becomes

$$(3.25) \quad s = \frac{\partial y}{\partial x}(5) + \frac{\partial y}{\partial t}(5) \left[\frac{t_1 - t_0}{h_i} \right] + \frac{h_i}{8} \left[\frac{\partial^2 y}{\partial x^2}(q_1) - \frac{\partial^2 y}{\partial x^2}(q_2) \right] \\ + \frac{(t_1 - t_0)}{4} \left[\frac{\partial^2 y}{\partial x \partial t}(q_1) - \frac{\partial^2 y}{\partial x \partial t}(q_2) \right] \\ + \frac{1}{2} \left[\frac{\partial^2 y}{\partial x \partial t}(q_1) + \frac{\partial^2 y}{\partial x \partial t}(q_2) \right] \cdot \left[\frac{t_1 + t_0}{2} - t_5 \right] \\ + \frac{1}{4} \left[\frac{\partial^2 y}{\partial t^2}(q_1) + \frac{\partial^2 y}{\partial t^2}(q_2) \right] (t_1 - t_0) \\ + \frac{1}{2} \left[\frac{\partial^2 y}{\partial t^2}(q_1) - \frac{\partial^2 y}{\partial t^2}(q_2) \right] \left[\frac{t_1 + t_0}{2} - t_5 \right].$$

Since $\frac{\partial y}{\partial x}(5) = \frac{B}{A}(5)$, what remains of (3.25) is the difference between $\frac{B}{A}(5)$ and the true slope s .

$$\left| \frac{\partial y}{\partial t}(5) \left(\frac{t_1 - t_0}{h_i} \right) \right| \leq K h_i^2$$

since $\frac{\partial y}{\partial t}$ is bounded and the method of computing the net assures us $t_1 - t_0$ is $O(h^3)$. Since $\frac{\partial^2 y}{\partial x^2} \in C^1$, if we re-expand

$$\frac{h_i}{8} \left[\frac{\partial^2 y}{\partial x^2}(q_1) - \frac{\partial^2 y}{\partial x^2}(q_2) \right]$$

about point 5, we see this term is $O(h^2)$.

Each of the remaining terms has a factor $\left[\frac{t_1 + t_0}{2} - t_5 \right]$. Since $y(x_5, t_1) = t_1$, $y(x_5, t_0) = t_0$, and $t_5 = y_5 = \frac{1}{2} [y(x_1, t_1) + y(x_0, t_0)]$,

we can write

$$(3.26) \quad \left[\frac{t_1 + t_0}{2} - t_5 \right] = \frac{1}{2} [y(x_5, t_1) + y(x_5, t_0) - y(x_1, t_1) - y(x_0, t_0)].$$

We expand each function on the right-hand side of (3.26) about

$(x_5, \frac{t_1 + t_0}{2})$. Recall y^* is the solution of $\bar{y}'' = B/A$ passing

through $(x_5, \frac{t_1 + t_0}{2})$. Thus the terms on the right hand side of equation (3.26) are as follows:

$$(3.27) \quad y(x_5, t_1) = y^*(x_5) + \frac{\partial y}{\partial t}(Q_3) \left(\frac{t_1 - t_0}{2} \right),$$

$$(3.28) \quad y(x_5, t_0) = y^*(x_5) + \frac{\partial y}{\partial t}(Q_4) \left(\frac{t_0 - t_1}{2} \right),$$

$$(3.29) \quad y(x_1, t_1) = y^*(x_5) + \frac{\partial y^*}{\partial x}(x_1 - x_5) + \frac{\partial y^*}{\partial t} \left(\frac{t_1 - t_0}{2} \right) \\ + \frac{1}{2} \frac{\partial^2 y}{\partial x^2}(Q_5)(x_1 - x_5)^2 + \frac{\partial^2 y}{\partial x \partial t}(Q_5)(x_1 - x_5) \left(\frac{t_1 - t_0}{2} \right) \\ + \frac{1}{2} \frac{\partial^2 y}{\partial t^2}(Q_5) \left(\frac{t_1 - t_0}{2} \right)^2,$$

and

$$(3.30) \quad y(x_0, t_0) = y^*(x_5) + \frac{\partial y^*}{\partial x}(x_5)(x_0 - x_5) + \frac{\partial y^*}{\partial t}(x_5) \left(\frac{t_0 - t_1}{2} \right) \\ + \frac{1}{2} \frac{\partial^2 y}{\partial x^2}(Q_6)(x_1 - x_5)^2 + \frac{\partial^2 y}{\partial x \partial t}(Q_6)(x_0 - x_5) \left(\frac{t_0 - t_1}{2} \right) \\ + \frac{1}{2} \frac{\partial^2 y}{\partial t^2}(Q_6) \left(\frac{t_1 - t_0}{2} \right)^2.$$

Therefore,

$$(\overset{\circ}{\varphi}, \overset{\circ}{\varphi})_R = (\bar{\varphi}, \bar{\varphi})_R = (\overset{\circ}{\varphi}, \overset{G}{\varphi})_R + (\bar{\varphi}, \overset{G}{\varphi})_R \quad (5.6)$$

$$= \frac{2^{\frac{1}{2}}}{\pi} (0_3) \left(\frac{2^{\frac{1}{2}}}{1 - \frac{1}{2}} \right) + \frac{2^{\frac{1}{2}}}{\pi} (0_4) \left(\frac{2^{\frac{1}{2}}}{2} \right) + \frac{2^{\frac{1}{2}}}{\pi} (0_5) \left(\frac{2^{\frac{1}{2}}}{2} \right) + \frac{2^{\frac{1}{2}}}{\pi} (0_6) \left(\frac{2^{\frac{1}{2}}}{2} \right) + \frac{2^{\frac{1}{2}}}{\pi} (0_7) \left(\frac{2^{\frac{1}{2}}}{2} \right) + \frac{2^{\frac{1}{2}}}{\pi} (0_8) \left(\frac{2^{\frac{1}{2}}}{2} \right)$$

These terms are all $O(h^2)$ and Theorem 3.2 is proved.

Therom 3.2b is more difficult since the explanations must be

Formaldehyde with the first and second formaldehyde

JO THE PRESIDENT OF THE UNITED STATES

Comment: The same theorem is true with regard to the line

3 to 0 and (E/A)(T).

MEMORANDUM

Let $A, B, C \in C_2^2$. Suppose A and C are bounded away from zero.

Suppose $u \in C^2$. Further assume that the method used to compute

the next section is 0.12).

$$(3.32) \quad 2A(5) \frac{h_1}{u(0)-u(1)} = -2 \left\{ (Au_x + Bu_y)(5) + h_1^2 \nabla(u, h_1) \right\}$$

where Γ_1 is a bounded function on $G + G_1$.

• Үлсн дунд нэгдүгээр байр эзлэх улсууд болон тэдний хөгжлийн үзүүлэлтүүд

Let $A, B, C \in C_1^*$. Suppose A and C are bounded away from zero.

Suppose $n \in \mathbb{C}^2$. Further assume that the method used to compute the

Then

$$(3.53) \quad \left\{ \begin{aligned} & 24(5)^{n-1} = -2 \\ & (An^x + Bn^y)(5) + n^1 n^1(n) \end{aligned} \right.$$

where w_1 has the property $\lim_{h \rightarrow 0} w_1(h) = 0$.

Theorem 3.3a depends upon the slope s of the line from 0 to 1, being equal to $(B/A)(5)$ plus an h^2 correction. Theorem 3.3b uses the fact the slope s is approximated by $(B/A)(5)$ with an error $h w(h)$, where $\lim_{h \rightarrow 0} w(h) = 0$. A proof is given for Theorem 3.3a

and a straightforward modification of this proof will suffice for proof of theorem 3.3b.

In this section on comparison of the difference operator $M_h u$ and Mu , we use $T_1(u, h)$ to denote bounded functions on $G + G'$, and $w_1(h)$ to denote functions with the property $\lim_{h \rightarrow 0} w_1(h) = 0$.

In order to establish the validity of Theorem 3.3a, we expand $u(1) - u(0)$ about point 5 and truncate at the third derivative.

$$\begin{aligned}
 (3.34) \quad u(0) - u(1) &= u(5) - u_x(5) \frac{h_1}{2} - u_y(5) \frac{h_1 s}{2} \\
 &+ u_{xx}(5) \frac{h_1^2}{8} + u_{xy}(5) \frac{h_1^2 s}{4} + u_{yy}(5) \frac{h_1^2 s^2}{8} \\
 &- \left\{ u(5) + u_x(5) \frac{h_1}{2} + u_y(5) \frac{h_1 s}{2} + u_{xx}(5) \frac{h_1^2}{8} \right. \\
 &\left. + u_{xy}(5) \frac{h_1^2 s}{4} + u_{yy}(5) \frac{h_1^2 s^2}{8} \right\} + \frac{h_1^3}{48} \cdot (\text{Function of third} \\
 &\text{derivatives of } u).
 \end{aligned}$$

The remainder term involving third derivatives is bounded on $G + G'$.

The second derivatives cancel so that we have

$$\left[-u_x(5) - u_y(5) \cdot s \right] h_1 + \frac{h_1^3}{48} \cdot (\text{remainder}).$$

Since $s = (B/A)(5) + h^2 \cdot (\text{bounded function})$, $A(5) \cdot s = B(5)$

$+ h^2$ (bounded function). Therefore,

$$(3.35) \quad -2A(5) \frac{u(0)-u(1)}{h_1} = -2 \left\{ (Au_x + Bu_y)(5) + h^2 T_1 \right\},$$

where T_1 is a bounded function of third derivatives of u and second derivatives of solutions of the equation $y' = B/A$, as asserted by the theorem.

Corollary to Theorem 3.3a:

Under the same hypothesis

$$(3.36) \quad 2a(7) \frac{u(3)-u(0)}{h_{1-1}} = -2 \left\{ (Au_x + Bu_y)(7) + h_{1-1}^2 T_2 \right\},$$

where T_2 is a bounded function on $G + G'$.

Under the hypothesis of theorem 3.3b

$$(3.37) \quad 2A(7) \frac{u(2)-u(0)}{h_{1-1}} = -2 \left\{ (Au_x + Bu_y)(7) + h_{1-1} w_2(z) \right\}.$$

Proof of Theorem 3.1a:

We write the remaining difference terms of $M_h u$, given by equations (3.3) through (3.9), in terms of derivatives and remainders.

$$(3.38) \quad \frac{u(2)-u(0)}{k_j} = u_y(0) + k_j^2 T_3,$$

where T_3 is a bounded function of third derivatives of u on $G + G'$.

$$(3.39) \quad \frac{u(0)-u(4)}{k_{j-1}} = u_y(8) + k_{j-1}^2 T_4,$$

where T_4 is a bounded function of third derivatives of u on $G + G'$.

In terms involving D and E , we estimate, using forward differences.

$$(3.40) \quad \frac{u(2)-u(0)}{k_j} = u_y(0) + k_j T_5,$$

and

$$(3.41) \quad \frac{u(1)-u(0)}{h_1} = u_x(0) + u_y(0)(B/A)(5) + h_1 T_5,$$

where T_5 and T_6 are bounded functions of the second derivatives of u on $G + G'$.

When we collect the above results we have

$$(3.42) \quad M_1 u = -\frac{2}{h_1+h_{j-1}} \left[(Au_x + Bu_y)(5) + h_1^2 T_1 \right] \left[1 + \frac{h_1}{k_j+h_{j-1}} \left[\frac{B}{A}(6) - \frac{B}{A}(8) \right] \right]$$

$$(3.43) \quad M_2 u = \frac{2}{h_1+h_{j-1}} \left[(Au_x + Bu_y)(7) + h_{j-1}^2 T_2 \right] \left[1 - \frac{h_1-h_{j-1}}{k_j+h_{j-1}} \left[\frac{B}{A}(6) - \frac{B}{A}(8) \right] \right]$$

$$(3.44) \quad M_3 u = -\frac{2}{k_j+k_{j-1}} \left[u_y(6) + k_j^2 T_3 \right] \left[c(6) - \frac{B^2}{A}(6) \right]$$

$$(3.45) \quad M_4 u = \frac{2}{k_j+k_{j-1}} \left[u_y(8) + k_{j-1}^2 T_4 \right] \left[c(8) - \frac{B^2}{A}(8) \right]$$

$$(3.46) \quad M_5 u = D(0) \left[-(B/A)(5)(u_y(0) + k_j T_5) + u_x(0) + (B/A)(5)u_y(0) \right] + h_1 T_6$$

and

$$(3.47) \quad M_6 u = B(0)(u_y(0) + k_j T_5).$$

Recall that T_1, T_2, T_3, T_4, T_5 , and T_6 , appearing in equations (3.42) through (3.47), are all bounded; hence, the terms containing these are all $O(h)$ on $G + G'$. We further modify $M_3 u$ and $M_4 u$.

$$(3.48) \quad M_3 u = u_y(6) \left[c(6) - \frac{B^2}{A}(6) \right] = (Bu_x + Cu_y)(6) - \frac{B}{A}(6)(Au_x + Bu_y)(6),$$

and

$$(3.49) \quad M_4 u = u_y(8) \left[c(8) - \frac{B^2}{A}(8) \right] = (Bu_x + Cu_y)(8) - \frac{B}{A}(8)(Au_x + Bu_y)(8).$$

We want to compare $Mu(0)$ with $(M_h u)(0)$. In order to make this comparison, we expand those terms not already of order h in $M_1 u$

in a Taylor's formula about the point 0. The needed expansions follow:

$$(3.50) \quad (Au_x + Bu_y)(5) = (Au_x + Bu_y)(0) + \frac{h_1}{2} \left[\frac{\partial}{\partial x}(Au_x + Bu_y)(0) + \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \frac{B}{A}(5) \right] + h_1^2 T_6$$

where T_6 is a bounded function of third derivatives of u , first and second derivatives of A and B , and the remainder of the estimate $(B/A)(5)$ for the slope of the line from 0 to 1. Likewise,

$$(3.51) \quad (Au_x + Bu_y)(7) = (Au_x + Bu_y)(0) - \frac{h_{i-1}}{2} \left[\frac{\partial}{\partial x}(Au_x + Bu_y)(0) - \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \frac{B}{A}(5) \right] + h_{i-1}^2 T_7,$$

where T_7 is a bounded function similar to T_6 formed by remainders of Taylor's series. Further,

$$(3.52) \quad \frac{B}{A}(6) = \frac{B}{A}(0) + \frac{\partial}{\partial y} \left(\frac{B}{A} \right) (0) \frac{k_j}{2} + k_j^2 T_8,$$

$$(3.53) \quad \frac{B}{A}(8) = \frac{B}{A}(0) - \frac{\partial}{\partial y} \left(\frac{B}{A} \right) (0) \frac{k_{j-1}}{2} + k_{j-1}^2 T_9,$$

$$(3.54) \quad (Bu_x + Cu_y)(6) = (Bu_x + Cu_y)(0) + \frac{\partial}{\partial y}(Bu_x + Cu_y)(0) \frac{k_j}{2} + k_j^2 T_{10}.$$

$$(3.55) \quad (Bu_x + Cu_y)(8) = (Bu_x + Cu_y)(0) - \frac{\partial}{\partial y}(Bu_x + Cu_y)(0) \frac{k_{j-1}}{2} + k_{j-1}^2 T_{11}.$$

$$(3.56) \quad (Au_x + Bu_y)(6) = (Au_x + Bu_y)(0) + \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \frac{k_j}{2} + k_j^2 T_{12}.$$

$$(3.57) \quad (Au_x + Bu_y)(8) = (Au_x + Bu_y)(0) + \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \frac{k_{j-1}}{2} + \frac{k_{j-1}^2}{2} T_{13},$$

and

$$(3.58) \quad \frac{B}{A}(5) = \frac{B}{A}(0) + h_1 T_{14},$$

where $T_8, T_9, T_{10}, T_{11}, T_{12}, T_{13}$, and T_{14} are remainders of Taylor's

series involving third derivatives of u , first and second derivatives of A , B , and C , and remainders from estimating the slope by (B/A)(5) and (B/A)(7). In each case, the function T_1 is bounded on $G + G^2$.

We now rewrite the parts of $M_1 u$, making the substitutions shown above.

$$(3.59) \quad M_1 u = -\frac{2}{h_1 + h_{1-1}} \left[(Au_x + Bu_y)(0) + \frac{\partial}{\partial x}(Au_x + Bu_y)(0) \frac{h_1}{2} \right. \\ \left. + \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \left(\frac{B}{A} \right)(0) \frac{h_1}{2} \right. \\ \left. + \frac{h_1}{2} \cdot (Au_x + Bu_y)(0) \cdot \frac{\partial}{\partial y} \left(\frac{B}{A} \right)(0) \right] + \frac{h_1^2 T_{15}}{h_1 + h_{1-1}}.$$

T_{15} is a bounded function which includes all those terms not already exhibited in (3.59) arising in the above approximation of $M_1 u$.

$$(3.60) \quad M_2 u = \frac{2}{h_1 + h_{1-1}} \left[(Au_x + Bu_y)(0) - \frac{\partial}{\partial x}(Au_x + Bu_y)(0) \frac{h_{1-1}}{2} \right. \\ \left. - \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \left(\frac{B}{A} \right)(0) \frac{h_{1-1}}{2} \right. \\ \left. - \frac{h_{1-1}}{2} (Au_x + Bu_y)(0) \cdot \frac{\partial}{\partial y} \left(\frac{B}{A} \right)(0) \right] + \frac{h_{1-1}^2 T_{16}}{h_1 + h_{1-1}}.$$

T_{16} represents the remainder terms similar to T_{15} .

$$(3.61) \quad M_3 u = -\frac{2}{k_j + k_{j-1}} \left[(Bu_x + Cu_y)(0) + \frac{\partial}{\partial y}(Bu_x + Cu_y)(0) \frac{k_j}{2} \right. \\ \left. - \frac{B}{A}(0)(Au_x + Bu_y)(0) - \frac{k_j}{2} \frac{\partial}{\partial y} \frac{B}{A}(0)(Au_x + Bu_y)(0) \right. \\ \left. - \frac{k_j}{2} \left(\frac{B}{A} \right)(0) \cdot \frac{\partial}{\partial y}(Au_x + Bu_y)(0) \right] - \frac{2k_j^2 T_{10}}{k_j + k_{j-1}} + \frac{k_j^2 T_{17}}{k_j + k_{j-1}},$$

where T_{17} represents remainder terms.

$$\begin{aligned}
(3.62) \quad M_4 u = & \frac{2}{k_j + k_{j-1}} \left[(Du_x + Cu_y)(0) - \frac{\partial}{\partial y} (Bu_x + Cu_y)(0) \frac{k_j-1}{2} \right. \\
& - \left(\frac{B}{A} \right) (0) (Au_x + Bu_y)(0) + \frac{k_j-1}{2} \frac{\partial}{\partial y} \left(\frac{B}{A} \right) (0) (Au_x + Bu_y) \\
& \left. + \frac{k_j}{2} \left(\frac{B}{A} \right) (0) \frac{\partial}{\partial y} (Au_x + Bu_y)(0) \right] + \frac{2k_j^2 T_{11}}{k_j + k_{j-1}} + \frac{k_j^2 T_{16}}{k_j + k_{j+1}}
\end{aligned}$$

where T_{18} represents the remainder terms.

$$(3.63) \quad M_5 u = (Du_x)(0) - D(0) \left[\frac{B}{A}(0) k_j T_6 - h_1 (T_5 + T_{14}) \right]$$

$$(3.64) \quad M_6 u = Bu_y(0) + (Du_y)(0) k_j T_6.$$

The T_i terms are bounded continuous functions representing terms not explicitly written. All such expressions containing T_i are order h expressions.

We now collect terms of $M_2 u$. The first terms on the right hand side of (3.59) and (3.60) cancel; the second terms can be combined and written as

$$(3.65) \quad - \frac{\partial}{\partial x} (Au_x + Bu_y)(0),$$

and the third terms reduce to

$$(3.66) \quad - \frac{\partial}{\partial y} (Au_x + Bu_y)(0) \left(\frac{B}{A} \right) (0).$$

The first terms on the right hand side of (3.61) and (3.62) cancel, and the second terms become

$$(3.67) \quad - \frac{\partial}{\partial y} (Bu_x + Cu_y)(0).$$

The third terms cancel; the fourth terms become

$$(3.68) \quad \frac{\partial}{\partial y} \left(\frac{B}{A} \right) (0) \cdot (Au_x + Bu_y)(0)$$

which cancels with the fourth terms of (3.59) and (3.60); the fifth terms combine to give

$$(3.69) \quad \left(\frac{B}{T}\right)(0)(Au_x + Bu_y)_y(0)$$

which, in turn, cancels with (3.66). All terms containing T_1 are of order h ; hence, we have

$$(3.70) \quad \begin{aligned} L_h u(0) = & -(Au_x + Bu_y)_x(0) - (Bu_x + Cu_y)_y(0) \\ & + (Du_x)(0) + (Eu_y)(0) + \text{terms of order } h. \end{aligned}$$

This completes the proof of Theorem 3.1a.

EXISTENCE OF THE SOLUTION OF THE DIFFERENCE EQUATION
AND CONVERGENCE TO THE SOLUTION OF THE DIFFERENTIAL EQUATION

Proofs of existence of solutions of the difference scheme and convergence to the correct solution of the differential equation have been done for a quite general Dirichlet problem in the literature. See Motzkin and Wasow (12), Bers (1), Parter (13), and Parter and Greenspan (9). Some of the above authors proved convergence for a different class of differential equations, and some for a restricted class where generalization was possible. Also convergence of a class of difference operators to a more restricted differential operator was done.

Convergence proofs all depend upon proving a maximum principle for the difference equation similar to that for elliptic differential equations. A sufficient condition for a maximum principle is that a difference scheme be of "positive type". This is a property our difference operator possesses and is defined later.

The convergence proofs presented here are adaptations of those given in Parter (13) in his treatment of the Laplacian. Further, most of Parter's proofs are either left to the reader or referenced to an earlier work of Bers (1) concerned with a special approximation to the Laplacian. We supply proofs to the theorems in the context of the difference equations of this paper.

Now we prove a maximum principle for the class of difference operators of positive type, a class which includes our difference operator. From this, the existence and uniqueness of the solution of the difference equations follow.

Let $H(h)$ denote the set of all real valued functions on the net region $G(h) + G'(h)$.

Definition 4.1:

Let M_h be a linear operator mapping $H(h)$ into $H(h)$ which has the following form:

$$(4.1) \quad (M_h u)(P) = \sigma(P, P)u(P) - \sum_{Q \in N(P)} \sigma(P, Q)u(Q), \quad P \in G(h).$$

We say M_h is of positive type if

$$(4.2) \quad \begin{aligned} a) \quad & \sigma(P, Q) > 0, \quad Q \in N(P) + \{P\}, \\ b) \quad & \sigma(P, P) - \sum_{Q \in N(P)} \sigma(P, Q) = 0. \end{aligned}$$

Note that our difference operator, the right hand side of equation (2.58), satisfies these conditions.

Theorem 4.1 (Maximum Principle):

Let $F(P)$ be a non-negative function defined on $G(h)$. Let $G(h) + G'(h)$ be connected. Let $(M_h u)(P)$ of positive type $\in H(h)$. If

$$(4.3) \quad (M_h u)(P) + F(P)u(P) \leq 0, \quad P \in G(h),$$

then

$$(4.4) \quad u(P) \leq \max \left\{ 0, \max_{Q \in G'(h)} u(Q) \right\}.$$

Moreover, if the same hypothesis holds and $u(P)$ assumes a non-negative maximum at an interior point, then $u(P)$ is constant.

Proof:

Deny the first conclusion.

Case 1. Assume $0 = \max \left\{ 0, \max_{Q \in G'(h)} u(Q) \right\}$, then $u(P) > 0$ at some

interior point. Choose from the set of points $P \in G(h)$

where $u(P)$ is a maximum, a point which has at least one neighbor Q such that $u(Q) < \max u(P)$.

From the hypothesis

$$\sigma(P, P)u(P) - \sum_{Q \in N(P)} \sigma(P, Q)u(Q) + F(P)u(P) \leq 0.$$

$$\begin{aligned} u(P) &\leq \frac{1}{\sigma(P, P) + F(P)} \sum_{Q \in N(P)} \sigma(P, Q)u(Q) \\ &< \frac{1}{\sigma(P, P) + F(P)} \sum_{Q \in N(P)} \sigma(P, Q)u(P), \end{aligned}$$

since $u(Q) < u(P)$ for at least one $Q \in N(P)$. Since

$$\sum_{Q \in N(P)} \sigma(P, Q) = \sigma(P, P) \text{ and } F(P) \geq 0, \text{ we have a contra-}$$

dition.

Case 2. $\max u(Q) = \max \left\{ 0, \max u(Q) \text{ for } Q \text{ in } G(h) \right\}.$

Again it is possible to choose from the set of points where $u(P)$ is maximum, a point which has at least one neighbor Q such that $u(Q) < u(P)$. Then the inequalities of Case 1 give the same contradiction. Therefore, the first conclusion holds.

Deny the second conclusion. Suppose $u(P)$ assumes a non-negative maximum in G , and $u(P)$ is not constant. Then the proof is again as Case 1 above. This new hypothesis allows us to choose a point P where at least one neighboring value $u(Q)$ is less than $u(P)$. Therefore, if u has a positive maximum in $G(h)$, then u is constant.

Corollary 1 (Minimum Principle):

If

$$(4.5) \quad (M_h u)(P) + F(P)u(P) \geq 0, \quad P \text{ in } G(h),$$

then

$$(4.6) \quad u(P) \geq \min \left\{ 0, \min_{Q \in G'(h)} u(Q) \right\}.$$

Moreover, for the same hypothesis, and if $u(P)$ assumes a minimum at an interior point P in $G(h)$ and this minimum is less than 0, then $u(P)$ is a constant.

Note that these are analogs of theorems for elliptic differential operators.

Corollary 2:

If

$$(4.7) \quad (M_{\bar{h}}u)(P) + F(P)u(P) = 0, \quad P \in G(h),$$

then

$$(4.8) \quad |u| \leq \max |u| \quad \text{on } G'(h).$$

Theorem 4.2:

Let $S(P)$ be any function defined on $G(h)$ and let $G(P)$ be any function defined on $G'(h)$. Then there exists a unique function $u(P) \in H(h)$ which satisfies

$$(4.9) \quad (M_{\bar{h}}u)(P) + F(P)u(P) = S(P), \quad P \in G(h),$$

where

$$(4.10) \quad u(P) = G(P), \quad P \in G'(h).$$

Proof:

This is a linear system of equations in as many unknowns as equations, one for each point $P \in G(h)$. The system

$$(4.11) \quad (M_{\bar{h}}u)(P) + F(P)u(P) = 0, \quad P \in G(h),$$

where

$$(4.12) \quad u(P) = 0, \quad P \in G'(h),$$

has only the trivial solution from the maximum principle; hence, the non-homogeneous system has a unique solution.

Since our difference equations satisfy all the needed hypothesis, we can conclude that they have a unique solution.

We have not yet shown convergence of the difference equation to the differential equation as $h \rightarrow 0$. To facilitate this, we first prove some convergence theorems and then show that our difference operator does satisfy the hypothesis of these theorems.

Assume we have a finite analog of the Dirichlet problem of the form

$$(4.13) \quad (M_{h(s)}u)(P) + F(P)u(P) = S(P), \quad P \in G(h(s))$$

$$(4.14) \quad u(P) = g(P), \quad P \in G'(h(s)).$$

Definition 4.2:

We say a sequence of nets $G(h(s)) + G'(h(s))$ "converge" to $G + G'$ in the sense:

- i) $h(s) \rightarrow 0$ as s
- ii) $G'(h(s)) \subset G'$.
- iii) g defined on $G'(h(s))$ is a subset of g defined on G' .

Some difficulty arises here in using the cited literature for convergence since they form sequences of nets by successive refinement so that if $P \in G(h(s))$, then $P \in G(h(s'))$ for all $s' \geq s$. The method of generating the net by numerical solution of ordinary differential equations excludes the possibility of net refinement.

Definition 4.3:

We say that the functions $u_s(P_s) \in G(h(s))$ converge uniformly

to $u(P_s) \in G$ as $s \rightarrow \infty$ if for $\epsilon > 0$ there is s_0 such that

$$|u_s(P_s) - u(P_s)| < \epsilon$$

for any $P_s \in G(h(s))$ when $s \geq s_0$.

Here the point set $\{P_s\}$ changes for changing s . The functions u_s change and are in general not comparable. However, $u(P_s)$ is defined for all $P_s \in G$ and so is always comparable to the functions $u_s(P_s)$.

Definition 4.4:

Let $M_{h(s)}$ be a linear operator of the form

$$\sigma(P, P)u(P) - \sum_{Q \in N(P)} \sigma(P, Q)u(Q)$$

of positive type mapping $H(h(s))$ into $H(h(s))$. We say the sequence

$\{M_{h(s)}\}$ is a "uniformly consistent" sequence of approximations

to the differential operator M if for every function $u(P)$ which is defined and twice differentiable on $G + G'$ we have

$$(4.15) \quad (M_{h(s)}u)(P) = (Mu)(P) + T(u, P_s, h(s)).$$

$T(u, P_s, h(s))$, the truncation error, is bounded in absolute value

by $j(u, h(s))$, a positive continuous function for $h \geq 0$, with

$$j(u, 0) = 0.$$

Theorem 4.3:

Let L be a uniformly elliptic operator of the form (1.2). Let $u \in C^2$. Then

$$(4.16) \quad |u(P)| \leq \max_{Q \text{ on } G'} |u(Q)| + K \max_{Q \text{ in } G} |(Mu)(P) + F(P)u(P)|$$

The proof is given in Courant and Hilbert (4, Vol. II, pp. 329-330).

For convergence of the difference operator, we need a discrete analog of this theorem.

Theorem 4.4:

Let $\{G(h(s)) + G'(h(s))\}$ be a sequence of nets converging to $G + G'$. Let $M_{h(s)}$ be a sequence of uniformly consistent approximations to the differential operator M , defined on the sequence of nets $\{G(h(s)) + G'(h(s))\}$. Let $F(x,y) \geq 0$ be defined on $\{G(h(s)) + G'(h(s))\}$. Then there are constants K and s_0 such that

$$(4.17) \quad |u| \leq \max_{\text{on } G'(h(s))} |u| + K \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu|$$

for all functions $u \in C^2$ on $G + G'$ and for all nets of the $\{h(s)\}$ such that $s \geq s_0$.

Proof:

Let a be a positive number to be specified later. Let x be the first coordinate of the point P in G . Let z be a number larger than any x associated with the set $G + G'$. There is always a number z since $G + G'$ is bounded. Let $u = e^{ax}$, so for each point P in G we evaluate at the first coordinate of P .

For $u = e^{ax}$

$$(4.18) \quad Mu(P) = M(e^{ax}) = (-a^2 A(P) - a(A_x(P) - D(P)))e^{ax},$$

where M is the differential operator (1.1).

We will be using the function $(e^{az} - e^{ax})$. Since $a > 0$ and $z > x$, $(e^{az} - e^{ax}) > 0$. We can assume without loss of generality that $x \geq 0$, i.e. that the region $G + G'$ is in the right-half plane. If the region $G + G'$ does not satisfy this condition, we would use the alternate function,

$$e^{a(z-y)} - e^{a(x-y)}.$$

y is the minimum of all x associated with $G + G'$. The proof would proceed in exactly the same way with this function as with $(e^{az} - e^{ax})$.

Returning to the assumption $G + G'$ is in the right-half plane, consider the auxiliary function

$$(4.19) \quad w(P) = u(P) - \max_{\text{in } G(h(s))} \left[M_{h(s)} u + Fu \right] \cdot (e^{az} - e^{ax}).$$

Then

$$(4.20) \quad \begin{aligned} (M_{h(s)} w)(P) + F(P)w(P) \\ = (M_{h(s)} u)(P) + F(P)u(P) - \max_{\text{on } G(h(s))} \left[M_{h(s)} u + Fu \right] \cdot M_{h(s)} (e^{az} - e^{ax}). \end{aligned}$$

But

$$(4.21) \quad M_{h(s)} u = Lu + F(u, P_s, h(s))$$

for any function of class C^2 . Therefore,

$$(4.22) \quad M_{h(s)} u (e^{az} - e^{ax}) = L(e^{az} - e^{ax}) + F(e^{az} - e^{ax}, h(s)).$$

Then

$$(4.23) \quad \begin{aligned} M_{h(s)} w + Fw \leq (M_{h(s)} u)(P) + F(P)u(P) \\ - \max_{\text{in } G(h(s))} \left[M_{h(s)} u + Fu \right] \left[e^{ax} \left\{ a^2 A + a(A_X - D) \right\} + F(e^{az} - e^{ax}, h(s)) \right]. \end{aligned}$$

$$\text{Choose } a \text{ such that } e^{ax} \left\{ a^2 A + a(A_X - D) \right\} \geq 2.$$

For this choice of a , it is possible to choose s large enough,

i.e. $h(s)$ small enough, so that

$$(4.24) \quad \left| F(e^{az} - e^{ax}, h(s)) \right| \leq j(e^{az} - e^{ax}, h(s)) < 1.$$

Then

$$(4.25) \quad M_{h(s)} w + Fw \leq$$

$$\begin{aligned}
& (M_{h(s)}u)(P) + F(P)u(P) - 2 \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \\
& = \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}, h(s)) \\
& \leq (M_{h(s)}u)(P) + F(P)u(P) - \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \\
& = \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (1 + (e^{ax} - e^{bx}, h(s))) \leq 0.
\end{aligned}$$

Therefore, we can apply the maximum principle to $M_{h(s)}w + Fw$.

$$w \leq \max \left\{ 0, \max w \text{ on } G^* \right\}.$$

If $w \leq 0$, then

$$u - \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}) \leq 0;$$

hence,

$$(4.26) \quad u \leq \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}).$$

If $w \leq \max w$ on G^* , then

$$\begin{aligned}
& u - \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}) \\
& \leq \max \left\{ u - \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}) \right\} \\
& \quad \text{on } G^*(h(s)) \text{ in } G(h(s))
\end{aligned}$$

Therefore,

$$(4.27) \quad u \leq \max_{\text{on } G^*(h(s))} |u| + \max_{\text{on } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{ax} - e^{bx}).$$

Let $K = \max (e^{ax} - e^{bx})$ on $G + G^*$. Then, combining the results of (4.26) and (4.27), we have

$$(4.28) \quad u \leq \max_{\text{on } G^*(h(s))} |u| + K \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu|$$

To obtain a lower bound on u of similar type, we apply the minimum principle to the function

$$(4.29) \quad w = u + \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu| \cdot (e^{sz} - e^{zh}).$$

Following nearly identical steps as in the above proof, we get a lower bound

$$(4.30) \quad u(P) \geq - \max_{\text{on } G^*(h(s))} |u| - K \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu|.$$

Combining the two inequalities (4.28) and (4.30), we obtain the desired result:

$$(4.31) \quad |u(P)| \leq \max_{\text{on } G^*(h(s))} |u| + K \max_{\text{in } G(h(s))} |M_{h(s)}u + Fu|$$

for all $s \geq s_0$.

This inequality enables us to prove convergence of a sequence of solutions of the difference equation to the solution of the differential equation when the sequence of difference operators is a uniformly consistent approximation to the differential operator.

Theorem 4.5:

Let $\{G(h(s)) + G^*(h(s))\}$ be a sequence of sets converging to $G + G^*$. Let $\{M_{h(s)}\}$ be a uniformly consistent sequence of approximations to the operator M on $G(h(s)) + G^*(h(s))$. Let U_s be a solution of the difference equation

$$(4.9) \quad (M_{h(s)}U)(P) + F(P)U_s(P) = S(P) \text{ in } G(h(s))$$

with the boundary condition

$$(4.10) \quad U_s(P) = G(P) \text{ on } G'(h(s)).$$

Let u satisfy the differential equation

$$(1.3) \quad (Mu)(P) + F(P)u(P) = S(P) \text{ in } G,$$

with boundary condition

$$(1.4) \quad u(P) = G(P) \text{ on } G'.$$

Then U_s converges uniformly to u in the sense of definition 4.3.

Since, as s approaches infinity, the net spacing approaches zero, and the number of points becomes infinite, we have a good approximation to the actual value of u at any point in $G + G'$ by choosing s large enough due to the continuity of u .

Proof:

Consider the function $U_s - u$.

$$\begin{aligned} & (M_{h(s)}(U_s - u))(P_s) + F(P_s)(U_s - u)(P_s) \\ &= (M_{h(s)}U_s)(P_s) - (Mu)(P_s) - T(u, P_s, h(s)) + F(P_s)(U_s - u)(P_s) \\ &= (M_{h(s)}U_s)(P_s) + F(P_s)U_s(P_s) - ((Mu)(P_s) + F(P_s)u(P_s)) \\ & \quad - T(u, P_s, h(s)). \end{aligned}$$

Therefore, since

$$\begin{aligned} & (M_{h(s)}U_s)(P_s) + F(P_s)U_s(P_s) = S(P_s), \text{ and} \\ & (Mu)(P_s) + F(P_s)u(P_s) = S(P_s), \end{aligned}$$

then

$$(4.32) \quad (M_{h(s)}(U_s - u))(P_s) + F(P_s)(U_s - u)(P_s) = -T(u, P_s, h(s)).$$

Using the inequality (4.31) on the function $(U_s - u)$,

$$(4.33) \quad |(U_s - u)(P_s)| \leq \max_{\text{on } G'(h(s))} |(U_s - u)(P_s)| +$$

truncation error is $O(1)$, the truncation error still depends upon
 Therefore, this value $K^* < \infty$ is guaranteed. Similarly, when the
 slope of the line joining net points, all on a bounded region $G + G^*$,
 the differential equation, and a uniformly bounded estimate of the
 derivatives of u , continuous derivatives of the coefficients of
 this is true since the truncation error depends upon continuous
 error $|u^n - u| \leq K^n$ for any n is smaller than a given ϵ_0 .
 function u , there is a constant $K^* > 0$ such that the truncation
 error is $O(h)$. That is, for any given
 and from the second set the convergence was only $O(h)$. In the
 initial equation. From one we obtained $O(h)$ convergence of u^n to u ,
 We have two sets of conditions on the coefficients of the differ-

value problem.
 So the difference equations have a unique solution for the boundary
 sufficiently small. Therefore, we can apply the maximum principle
 We have already observed that $u^n(c)$ is of positive type when n is
 does satisfy the necessary hypotheses to apply these theorems.
 We must now show that indeed the difference scheme in this

This completes the proof of Theorem 4.5.
 $\epsilon > 0$, there is an ϵ_0 such that if $\epsilon \leq \epsilon_0$, then $\delta(u, u^n(c)) < \epsilon$.
 and (4.7). $\delta(u, u^n(c))$ then depends only upon ϵ . For any pre-assigned
 We have a fixed u , the solution to the differential equation (1.3)

$$(4.7) \quad |u^n(c) - u(c)| \leq K \cdot |u^n(c) - u(c)| \leq K \delta(u, u^n(c))$$

where $u^n(c) - u(c) = 0$ on $G + G^*$, hence

$$|u^n(c) - u(c)| \leq K \max_{G + G^*} |u^n(c) - u(c)| + K \delta(u, u^n(c))$$

continuous derivatives of u and the coefficients of the differential equation. Therefore, function $j(u, h(s))$ can be defined for fixed u as follows. For each s , let $j(u, h(s))$ be the supremum of the absolute value of all truncation errors for step size $\leq h(s)$. The uniform continuity of the variables in the truncation error makes $j(u, h(s))$ continuous, and j can be made arbitrarily small by choosing h small enough. We already showed the truncation error vanished as $h \rightarrow 0$. Hence, the theorems apply to our difference operator. Therefore, the difference equation converges to the differential equation.

COMPARISON OF METHODS AND RESULTS

As an example, consider the boundary value problem involving the differential equation

$$(5.1) \quad u_{xx} + \frac{4}{3} xy u_{xy} + \frac{1}{2} xy u_{yy} = 0.$$

Greenspan and Jain (8) use this example on the unit square with corners $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$ using a nine point difference formula which gives an operator of positive type. The net used is rectangular and the difference equations and the net are "natural" to the region in the sense that no interpolation is needed to determine boundary points. The values on the boundary are generated by evaluation of

$$(5.2) \quad u = xy + \frac{4}{3} y(1 - y)$$

on the boundary for solution of the difference equations. This is also a solution of (5.1). The discretization error is a function of the third derivatives of the solution multiplied by step size, i.e. $O(h)$. Since the third derivatives vanish, the difference equation and the differential equation have the same solution. The solutions obtained by Greenspan and Jain match the accuracy of the machine.

The method we propose requires extra computation, the solution of $dy/dx = B/A$. However, this initial calculation reduces the number of neighbors from eight to four for each interior point and results in simpler difference equations, which, in turn, implies less machine operations per iteration. We use a "natural" region as shown in figure 7, page 69, where no interpolation is needed on

the boundary. The discretization procedure used approximates the boundary of the net region by straight lines. Further, the slope of the lines connecting net points are approximated by the value of B/A at the midpoint of the line connecting the points. Therefore, we cannot expect, and do not get, exact values when solving the difference equations on this problem. Except in special cases, neither our method nor the method of Greenspan and Jain will give an exact solution.

Some other comparisons between the methods can be made. Our method includes terms involving first partial derivatives. Their inclusion requires that the net be "sufficiently small" to insure that the differential operator is of positive type. Greenspan and Jain (8) require different formulation of the difference equations depending on the sign of the mixed derivative term of (1.1); our method does not. Even though both methods require that the equation be uniformly elliptic in order to insure convergence, both the example of Greenspan and Jain and ours use a region where (1.3) is parabolic on two sides of the boundary. Both methods give good results in this example.

Starting with the standard form, $u_{xx} + 2Bu_{xy} + Cu_{yy}$, for the elliptic operator, Greenspan and Jain use uniform step size in both the h direction and in the k direction, i.e. $h_i = h_{i-1}$ and $k_j = k_{j-1}$. With this restriction in net spacing, they require the ratio k/h to satisfy

$$(5.3) \quad B^2(x,y) \leq \frac{k}{h} |B(x,y)| < C(x,y)$$

at each point to ensure an operator of positive type. To obtain a theorem in the large, they choose a condition to ensure (5.3) as follows: If

$$(5.4) \quad \sup |B(x,y)| \leq \inf \frac{C(x,y)}{|B(x,y)|},$$

then the selection of h and k so that

$$(5.5) \quad \sup |B(x,y)| \leq \frac{k}{h} \leq \inf \frac{C(x,y)}{|B(x,y)|}$$

will satisfy (5.3). These restrictions exclude solving

$$(5.6) \quad u_{xx} + 2x u_{xy} + (x + 1/9) u_{yy} = 0,$$

since (5.4) is not satisfied. Our method has no such restrictions.

The restriction that the equation be uniformly elliptic is sufficient to guarantee that the difference equation be of positive type. This in turn gives us a maximum principle so that uniqueness of the solution of the difference operator and convergence to the differential operator is assured. Further, we are assured that an iterative method, such as GaussSeidel, will solve the difference equations.

Our method has no restrictions requiring uniformity of the h_i step size. The initial step in the y direction can be chosen. Thereafter, in general, the method of computing the net precludes the k_j from being equal. Further, the freedom in choice of step size allows interpolation on the boundary without changing the order of the method.

Our program for approximating the solution of the boundary

value problem is written in two parts. First, the equation $dy/dx = B/A$ is solved numerically to obtain net points. A machine plot showing net points and curves $\xi = \text{constant}$ is made. An example appears in figure 7, page 69.

Since the generated net may not coincide with the boundary of the region $G + G'$, interpolation on the boundary, if needed, is done at this stage. We have a region in which they coincide for the example.

The theory developed in this paper solves the equation

$$(5.7) \quad - (Au_x + Bu_y)_x - (Bu_x + Cu_y)_y + Du_x + Eu_y + Fu = S.$$

We first put the equation

$$(5.1) \quad u_{xx} + \frac{4}{3} xy u_{xy} + \frac{1}{2} xy u_{yy} = 0$$

in that form. We then see that $A = 1$, $B = \frac{2}{3} xy$, and $C = \frac{1}{2} xy$.

Then the first two terms of the left-hand side of equation (5.7) are

$$(5.8) \quad - (u_x + \frac{2}{3} xy u_y)_x - (\frac{2}{3} xy u_x + \frac{1}{2} xy u_y)_y = \\ - (u_{xx} + \frac{4}{3} xy u_{xy} + \frac{1}{2} xy u_{yy}) - \frac{2}{3} y u_y - \frac{x}{2} u_y - \frac{2}{3} x u_x,$$

so

$$(5.9) \quad D = \frac{2}{3} x,$$

and

$$(5.10) \quad E = \frac{2}{3} y + \frac{x}{2}.$$

Next, the coordinates of the net, the neighbors of each interior point, the boundary values and the differential equation are used as input in the second program. A system of equations is generated and solved by Gauss Seidel iteration. No attempt is made to speed

up convergence of the iteration.

Figure 7, page 69, shows the net region with a spacing of $1/10$ in the x direction and in the y direction for the starting points of solutions of $dy/dx = 5/x$. This gave 45 equations to solve. The solution took 106 iterations. The entire program, including compilation, computation of the system of equations and solution of the system took less than 42 seconds. Results are shown on page 70. Since the exact solution of the differential equation is known, errors are shown. Figure 9, page 71, shows the net region with the step size cut in half. There are now 209 interior points in the net region, so there is a system of 209 equations to solve. The solution of the system took 369 iterations. Total running time for the program was two minutes and 43 seconds. The fourth order Runge Kutta method used to form the net is accurate enough in this case so that within machine accuracy we actually subdivided the net. Some typical results are shown on page 72. Figure 11, page 73, shows the net region with the step size cut in half again. This time there are 882 equations to solve. Further, there are 4361 non zero entries in the coefficient matrix for the system of equations. The program is written so that only non zero entries are stored. For this size problem, all the information about the coefficient matrix is stored in 15,000 locations versus the 810,000 locations needed if the whole matrix were stored. The solution of the system took 1600 iterations, accumulated on four different runs. Actual running time for the program was 43 minutes,

of which approximately two-thirds of the time was used in iterative solution of the system of equations.

A comparison of the accuracy for each step size at the same point is shown for several different points on page 75. Note that the error at the points is cut in half when the net spacing is cut in half. This is the expected asymptotic value, since the method is of order h . The points chosen show extremes in the errors for the region as well as intermediate results. Errors are smaller near the boundary of the region for any given net size.

BIBLIOGRAPHY

1. Bers, Lipman. On mildly nonlinear partial difference equations of elliptic type. National Bureau of Standards Journal of Research, 51: 229-236. 1953.
2. Bers, Lipman, Fritz John and Martin Schechter. Partial differential equations. In Lectures in applied mathematics. Volume 3. Interscience Publishers, New York, New York. 1964.
3. Coddington, Earl A. and Norman Levinson. Theory of ordinary differential equations. McGraw-Hill Book Company, Inc., New York, New York. 1955.
4. Courant, R. Partial differential equations. In Courant, R. and D. Hilbert. Methods of mathematical physics. Volume 2. Interscience Publishers, New York, New York. 1962.
5. Garabedian, P. R. Partial differential equations. John Wiley and Sons, Inc., New York, New York. 1964.
6. Greenspan, Donald. Note on difference approximations with non-negative coefficients for partial differential equations. University of Wisconsin Mathematics Research Center Technical Summary Report 323. 1962.
7. Greenspan, Donald. On the approximate solution of elliptic differential equations with mixed partial derivatives. University of Wisconsin Mathematics Research Center Technical Summary Report 209. 1960.
8. Greenspan, Donald and P. O. Jain. On non negative difference analogues of elliptic differential equations. University of Wisconsin Mathematics Research Center Technical Summary Report 490. 1964.
9. Greenspan, Donald and Seymour V. Parter. Mildly nonlinear elliptic partial differential equations and their numerical solution. II. University of Wisconsin Mathematics Research Center Technical Summary Report 474. 1964.
10. Henrici, Peter. Discrete variable methods in ordinary differential equations. John Wiley and Sons, Inc., New York, New York. 1962.
11. Mac Neal, R. H. An asymmetrical finite difference network. Quarterly of Applied Mathematics 11: 295-310. 1953.

12. Motzkin, T. S. and W. Wasow. On the approximation of linear elliptic differential equations by difference equations with positive coefficients. *Journal of Mathematics and Physics* 31: 253-259. 1953.
13. Parter, Seymour V. Mildly nonlinear elliptic partial differential equations and their numerical solution. I. University of Wisconsin Mathematics Research Center Technical Summary Report 470. 1964.
14. Varga, Richard S. *Matrix iterative analysis*. Prentice-Hall Inc., Englewood Cliffs, New Jersey. 1962.
15. Wachspress, Eugene L. *Iterative solution of elliptic systems and application to the neutron diffusion equations of reactor physics*. Prentice-Hall Inc., Englewood Cliffs, New Jersey. 1966.
16. Weinberger, H. P. *A first course in partial differential equations with complex variables and transform methods*. First edition. Blaisdell Publishing Company, New York, New York. 1965.

ACKNOWLEDGMENT

I am particularly indebted to Dr. Clair G. Boyle for his guidance and criticism during the preparation of this dissertation.

APPENDIX

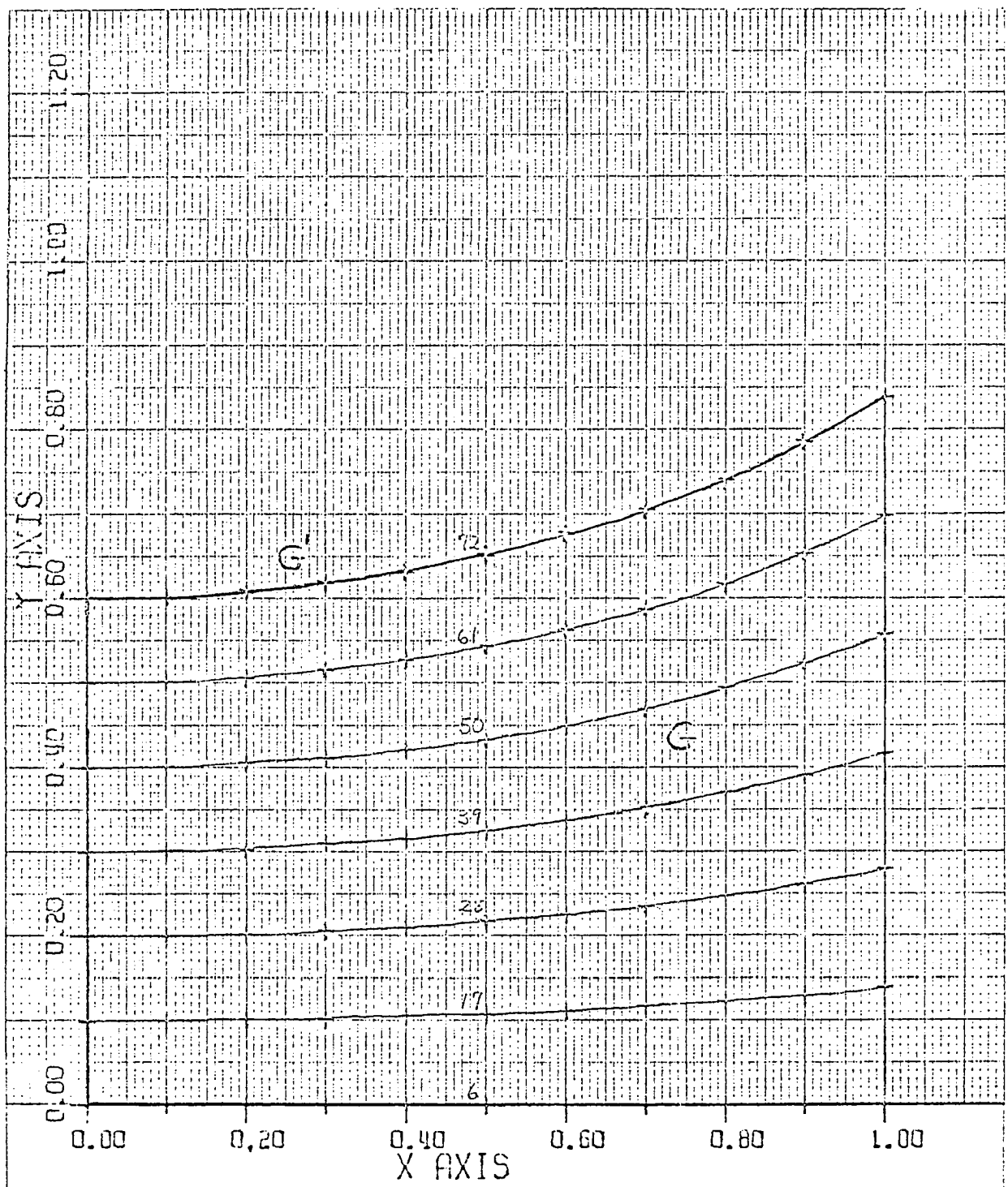
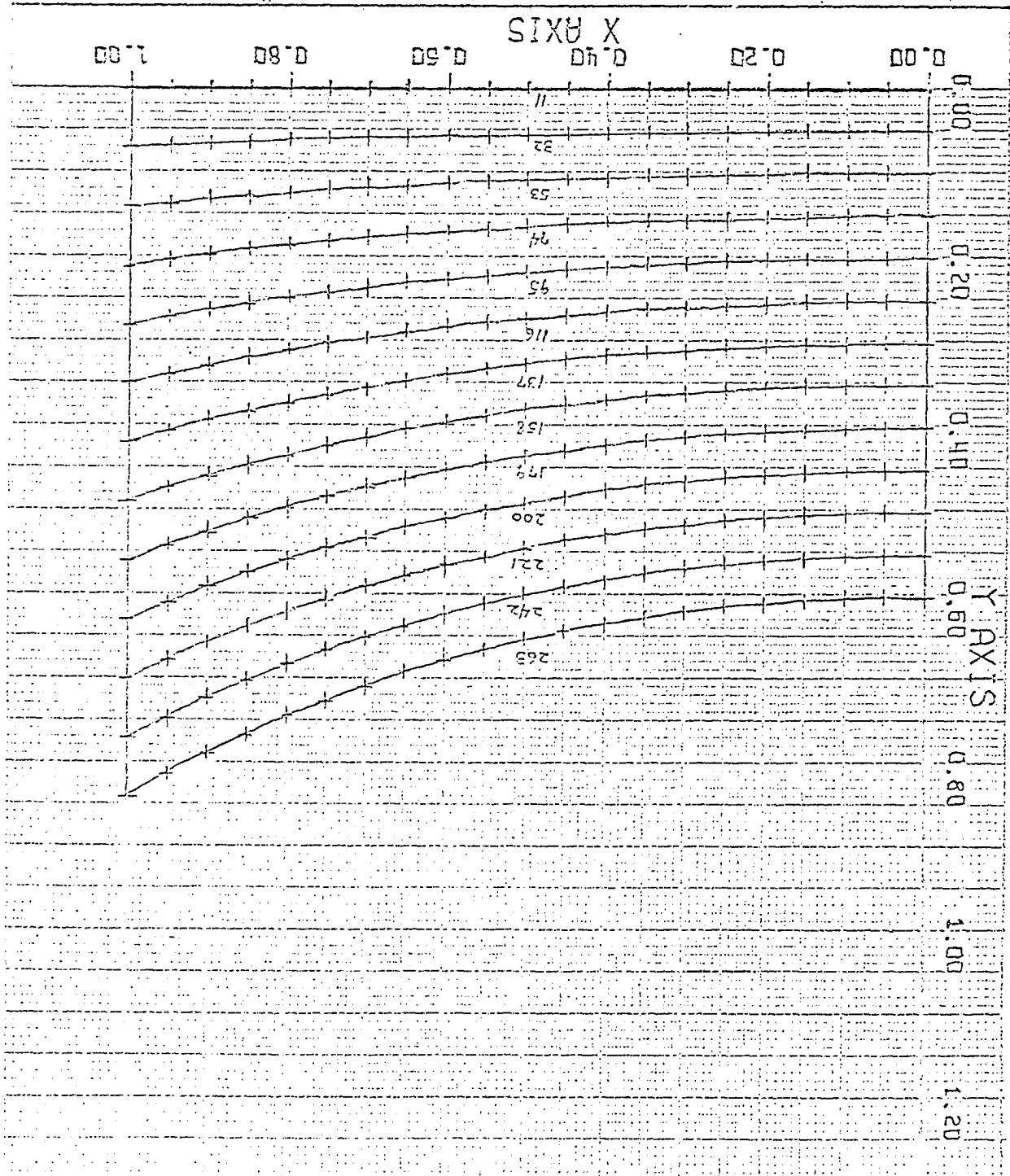


Figure 7. Region $G + G'$ with Net Size $h_1 = .1$

NET POINT	COMPUTED VALUE	CORRECT VALUE	ERROR	% ERROR
13	0.151807	0.130389	-0.001418	-1.0876%
14	0.144384	0.141698	-0.002687	-1.8962%
15	0.157809	0.154149	-0.003740	-2.4324%
16	0.172544	0.167994	-0.004550	-2.7084%
17	0.188545	0.183514	-0.005031	-2.7416%
18	0.206161	0.201032	-0.005129	-2.5514%
19	0.225697	0.220926	-0.004771	-2.1592%
20	0.247506	0.243634	-0.003871	-1.5890%
21	0.272902	0.269678	-0.003224	-0.9515%
24	0.235752	0.233934	-0.001818	-0.7773%
25	0.259408	0.256008	-0.003400	-1.3279%
26	0.284669	0.279084	-0.005585	-1.9733%
27	0.311936	0.306320	-0.005615	-1.8332%
28	0.341655	0.335525	-0.006129	-1.8262%
29	0.374327	0.368165	-0.006162	-1.6738%
30	0.410527	0.404882	-0.005644	-1.3941%
31	0.450912	0.446412	-0.004500	-1.0081%
32	0.496244	0.493596	-0.002648	-0.5365%
35	0.312784	0.310633	-0.002151	-0.6925%
36	0.346910	0.342931	-0.003979	-1.1603%
37	0.382929	0.377502	-0.005427	-1.4375%
38	0.421417	0.414978	-0.006439	-1.5517%
39	0.462993	0.456034	-0.006959	-1.5269%
40	0.508326	0.501398	-0.006929	-1.3919%
41	0.558155	0.551870	-0.006285	-1.1389%
42	0.613294	0.608332	-0.004962	-0.8157%
43	0.674645	0.671754	-0.002891	-0.4303%
46	0.362925	0.360487	-0.002438	-0.6762%
47	0.406930	0.402467	-0.004463	-1.1088%
48	0.452729	0.446705	-0.006024	-1.3486%
49	0.501047	0.493968	-0.007079	-1.4331%
50	0.552624	0.545039	-0.007585	-1.3917%
51	0.608229	0.600731	-0.007499	-1.2482%
52	0.668655	0.661899	-0.006765	-1.0221%
53	0.734717	0.729396	-0.005321	-0.7255%
54	0.807245	0.804153	-0.003092	-0.3845%
57	0.385953	0.383496	-0.002456	-0.6405%
58	0.433994	0.434615	-0.000621	-0.1432%
59	0.493350	0.487592	-0.005758	-1.1809%
60	0.549911	0.543289	-0.006622	-1.2199%
61	0.609536	0.602542	-0.006994	-1.1608%
62	0.673036	0.666164	-0.006873	-1.0317%
63	0.741155	0.734939	-0.006216	-0.8458%
64	0.814542	0.809602	-0.004940	-0.6101%
65	0.893708	0.890791	-0.002918	-0.3275%

Figure 8. Approximate Solution of $u_{xx} + \frac{4}{3}xy u_{xy} + \frac{xy}{2} u_{yy} = 0$
When $h_i = .1$

Figure 9. Region G + G' with Net Size $n_1 = .05$



NET POINT	COMPUTED VALUE	CORRECT VALUE	ERROR	% ERROR
41	0.148715	0.148153	-0.000562	-0.3793%
44	0.125458	0.125093	-0.000365	-0.2917%
45	0.131104	0.130389	-0.000715	-0.5482%
46	0.136861	0.135915	-0.000946	-0.7699%
47	0.143053	0.141698	-0.001356	-0.9562%
48	0.149405	0.147766	-0.001639	-1.1094%
49	0.156043	0.154149	-0.001894	-1.2285%
50	0.162996	0.160881	-0.002115	-1.3146%
51	0.170293	0.167904	-0.002299	-1.3687%
52	0.177969	0.175625	-0.002443	-1.3919%
53	0.186056	0.183514	-0.002542	-1.3854%
54	0.194594	0.192001	-0.002593	-1.3503%
55	0.203622	0.201032	-0.002590	-1.2882%
56	0.213185	0.210656	-0.002529	-1.2006%
57	0.223330	0.220925	-0.002405	-1.0887%
58	0.234110	0.231897	-0.002212	-0.9541%
59	0.245579	0.243634	-0.001945	-0.7985%
60	0.257800	0.256203	-0.001598	-0.6255%
61	0.270840	0.269677	-0.001162	-0.4310%
62	0.284770	0.284138	-0.000632	-0.2225%
171	0.361660	0.360487	-0.001173	-0.3254%
172	0.382928	0.381242	-0.001685	-0.4420%
173	0.404610	0.402467	-0.002143	-0.5325%
174	0.426800	0.424256	-0.002544	-0.5997%
175	0.449591	0.446705	-0.002887	-0.6462%
176	0.473078	0.469910	-0.003168	-0.6741%
177	0.497354	0.493968	-0.003386	-0.6856%
178	0.522517	0.518977	-0.003539	-0.6820%
179	0.548663	0.545039	-0.003625	-0.6650%
180	0.575395	0.572255	-0.003140	-0.6361%
181	0.602312	0.600731	-0.001582	-0.5983%
182	0.630019	0.630572	0.000552	0.5466%
183	0.665120	0.661889	-0.003231	-0.4882%
184	0.697723	0.694793	-0.002930	-0.4218%
185	0.731936	0.729396	-0.002540	-0.3483%
186	0.767868	0.765811	-0.002056	-0.2685%
187	0.805627	0.804152	-0.001475	-0.1834%
188	0.845322	0.844531	-0.000791	-0.0937%
191	0.353193	0.352568	-0.000625	-0.1772%
192	0.376543	0.375347	-0.001196	-0.3185%

Figure 10. Typical Approximations of $u_{xx} + \frac{4}{3}xy u_{xy} + \frac{xy}{2}u_{yy} = 0$
When $h_i = .05$

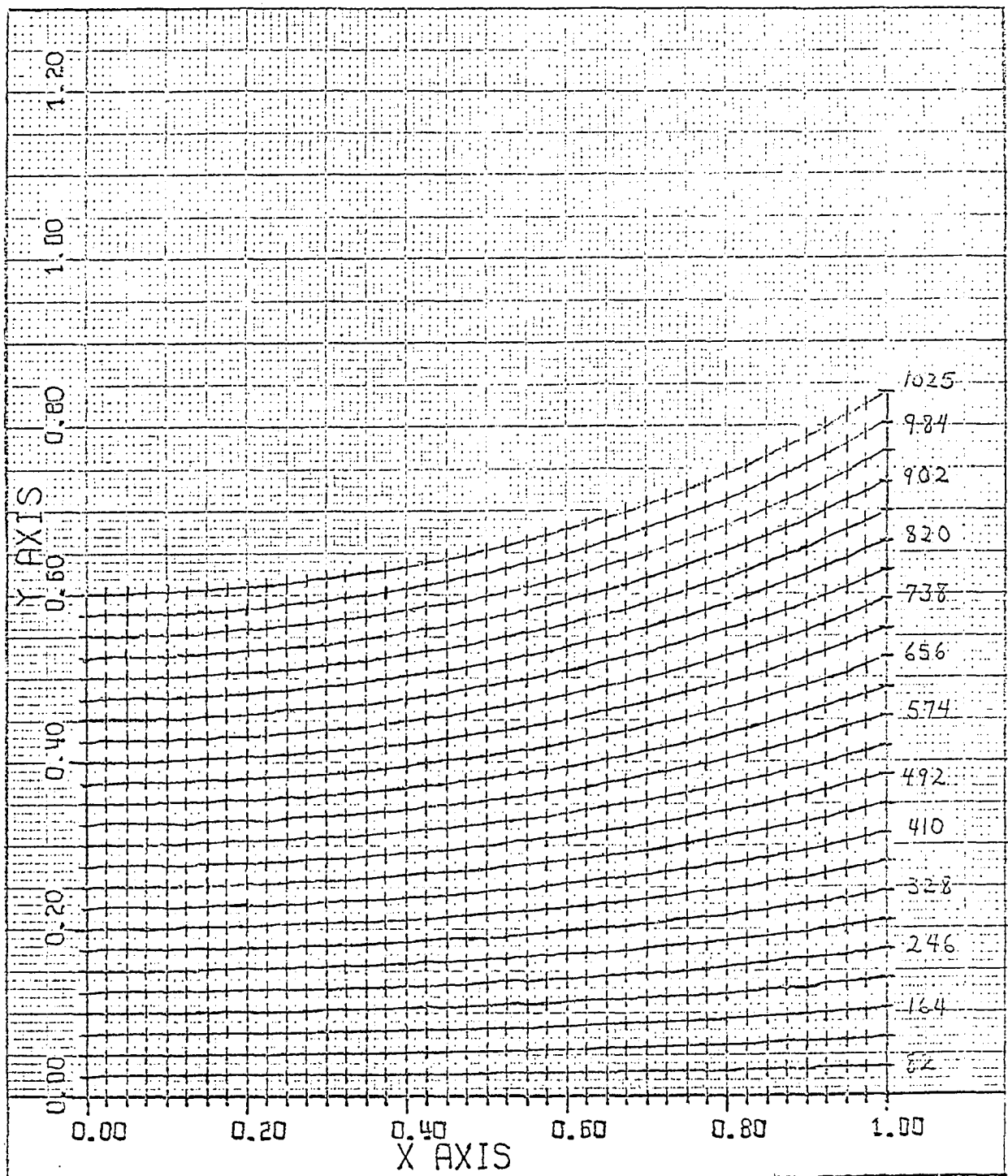


Figure 11. Region $G + G'$ with Net Size $h_1 = .025$

NET POINT	COMPUTED VALUE	CORRECT VALUE	ERROR	% ERROR
169	0.130745	0.130289	-0.000456	-0.34977
170	0.132861	0.133122	-0.000261	-0.19612
171	0.135436	0.135915	-0.000479	-0.35277
172	0.138372	0.138772	-0.000400	-0.28856
173	0.142373	0.141697	-0.000676	-0.47687
174	0.145442	0.144694	-0.000748	-0.51482
175	0.148582	0.147765	-0.000817	-0.55257
176	0.151798	0.150916	-0.000882	-0.58492
177	0.155094	0.154149	-0.000944	-0.61264
178	0.158471	0.157469	-0.000992	-0.63527
179	0.161926	0.160881	-0.001045	-0.65682
180	0.165491	0.164387	-0.001104	-0.67157
181	0.169141	0.167994	-0.001148	-0.68327
182	0.172891	0.171705	-0.001186	-0.69261
183	0.176744	0.175525	-0.001219	-0.69972
184	0.180706	0.179469	-0.001237	-0.70497
185	0.184782	0.183513	-0.001269	-0.70954
186	0.188976	0.187691	-0.001285	-0.71344
187	0.193294	0.192000	-0.001294	-0.71682
188	0.197741	0.196445	-0.001297	-0.71981
189	0.202323	0.201031	-0.001292	-0.72242
190	0.207046	0.205766	-0.001280	-0.72473
191	0.211916	0.210655	-0.001261	-0.72681
192	0.216939	0.215705	-0.001234	-0.72764
193	0.222123	0.220934	-0.001189	-0.72727
194	0.227473	0.226316	-0.001155	-0.72544
195	0.232998	0.231896	-0.001102	-0.72231
196	0.238705	0.237665	-0.001040	-0.71804
197	0.244601	0.243632	-0.000969	-0.71271
198	0.250695	0.249808	-0.000886	-0.70647
199	0.256996	0.256201	-0.000794	-0.70014
200	0.263512	0.262820	-0.000691	-0.69314
201	0.270253	0.269676	-0.000577	-0.68527
202	0.277229	0.276777	-0.000452	-0.67625
203	0.284450	0.284136	-0.000314	-0.66647
204	0.291926	0.291763	-0.000163	-0.65603
207	0.149383	0.148985	-0.000398	-0.66887
819	0.338470	0.338281	-0.000189	-0.60327
822	0.345985	0.345826	-0.000159	-0.64334
823	0.358645	0.358354	-0.000292	-0.68144
824	0.371328	0.370902	-0.000426	-0.71484
825	0.384049	0.383496	-0.000553	-0.74418
826	0.396822	0.396150	-0.000672	-0.76967
827	0.409661	0.408879	-0.000782	-0.79143
828	0.422581	0.421696	-0.000885	-0.80927
829	0.435594	0.434615	-0.000979	-0.82534

Figure 12. Typical Approximations of $u_{xx} + \frac{4}{3}xy u_{xy} + \frac{xy}{2} u_{yy} = 0$
 When $h_i = .025$

Step Size	Calculated Value	Error	% Error
<u>Point 38</u> $x = 0.500, y = 0.118754$			
.1	0.200081	-0.003129	-2.5514
.05	0.203622	-0.002590	-1.2682
.025	0.202323	-0.001292	-0.6427
Correct value	0.201032		
<u>Point 41</u> $x = 0.500, y = 0.130996$			
.1	0.272002	-0.002324	-0.8616
.05	0.270040	-0.001162	-0.4320
.025	0.270253	-0.000577	-0.2140
Correct value	0.269677		
<u>Point 39</u> $x = 0.500, y = 0.126071$			
.1	0.452593	-0.006959	-1.5260
.05	0.459398	-0.003365	-0.7370
.025	0.457683	-0.001650	-0.3618
Correct value	0.456033		
<u>Point 42</u> $x = 0.500, y = 0.171359$			
.1	0.615294	-0.004962	-0.8157
.05	0.610722	-0.002369	-0.3928
.025	0.609491	-0.001160	-0.1907
Correct value	0.608331		
<u>Point 48</u> $x = 0.500, y = 0.412182$			
.1	0.452729	-0.006024	-1.4603
.05	0.449591	-0.002667	-0.6462
.025	0.448116	-0.001411	-0.3160
Correct value	0.446705		
<u>Point 51</u> $x = 0.500, y = 0.450996$			
.1	0.605229	-0.007499	-1.2482
.05	0.604312	-0.003582	-0.5963
.025	0.602462	-0.001731	-0.2882
Correct value	0.600730		
<u>Point 57</u> $x = 0.100, y = 0.501690$			
.1	0.385553	-0.002456	-0.6105
.05	0.384637	-0.001141	-0.2975
.025	0.384049	-0.000553	-0.1441
Correct value	0.383469		

Figure 13. Comparison of Results for the Three Not Spacings at Some Typical Points